(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 April 2003 (24.04.2003)

PCT

(10) International Publication Number
WO 03/033688 A1

(51) International Patent Classification[7]: C12N 9/16, 15/55, C07K 16/40, C12Q 1/42, 1/68, G01N 33/573, A61K 38/46

(21) International Application Number: PCT/EP02/11473

(22) International Filing Date: 14 October 2002 (14.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/329,329 16 October 2001 (16.10.2001) US

(71) Applicant (*for all designated States except US*): **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): **LIOU, Jiing-Ren** [CN/US]; 10 Winslow Road, Belmont, MA 02478 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/033688 A1

(54) Title: REGULATION OF HUMAN RECEPTOR TYROSINE PHOSPHATASE

(57) Abstract: Reagents that regulate human receptor tyrosine phosphatase and reagents which bind to human receptor tyrosine phosphatase gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, diabetes, CNS disorders, obesity, COPD or cardiovascular disorders.

REGULATION OF HUMAN RECEPTOR TYROSINE PHOSPHATASE

FIELD OF THE INVENTION

The invention relates to the regulation of human receptor tyrosine phosphatase.

BACKGROUND OF THE INVENTION

Phosphorylation of proteins is a fundamental mechanism for regulating diverse cellular processes. U.S. Patent 5,952,212. While the majority of protein phosphorylation occurs at serine and threonine residues, phosphorylation at tyrosine residues is attracting a great deal of interest since the discovery that many oncogene products and growth factor receptors possess intrinsic protein tyrosine kinase activity. The importance of protein tyrosine phosphorylation in growth factor signal transduction, cell cycle progression and neoplastic transformation is now well established. Hunter *et al.*, *Ann. Rev. Biochem. 54*, 987-30, 1985; Ullrich *et al.*, Cell 61, 203-12, 1990; Nurse, *Nature 344*, 503-08, 1990; Cantley *et al.*, *Cell 64*, 281-302, 1991.

Biochemical studies have shown that phosphorylation on tyrosine residues of a variety of cellular proteins is a dynamic process involving competing phosphorylation and dephosphorylation reactions. The regulation of protein tyrosine phosphorylation is mediated by the reciprocal actions of protein tyrosine kinases (PTKases or PTKS) and protein tyrosine phosphatases (PTPs). The tyrosine phosphorylation reactions are catalyzed by PTKs. Tyrosine phosphorylated proteins can be specifically dephosphorylated through the action of PTPs. The level of protein tyrosine phosphorylation of intracellular substances is determined by the balance of PTK and PTP activities. Hunter, *Cell 58*, 1013-16, 1989.

REGULATION OF HUMAN RECEPTOR TYROSINE PHOSPHATASE

Protein Tyrosine Kinases

The protein tyrosine kinases (PTKS; ATP:protein-tyrosine O-phosphotransferase, EC 2.7.1.112) are a large family of proteins that includes many growth factor receptors and potential oncogenes. Hanks *et al.*, *Science 241*, 42-52, 1988. Many PTKs have been linked to initial signals required for induction of the cell cycle. Weaver *et al.*, *Mol. Cell. Biol. 11(9)*, 4415-22, 1991. PTKs comprise a discrete family of enzymes having common ancestry with, but major differences from, serine/threonine-specific protein kinases. Hanks *et al.*, *supra*. The mechanisms leading to changes in activity of PTKs are best understood in the case of receptor-type PTKs having a trans-membrane topology. Ullrich *et al.*, 1990, *supra*. The binding of specific ligands to the extracellular domain of members of receptor-type PTKs is thought to induce their oligomerization, leading to an increase in tyrosine kinase activity and activation of the signal transduction pathways. Ullrich *et al.*, 1990, *supra*. Deregulation of kinase activity through mutation or overexpression is a well-established mechanism for cell transformation. Hunter *et al.*, 1985; Ullrich *et al.*, 1990.

Protein Tyrosine Phosphatases

The protein phosphatases are composed of at least two separate and distinct families: the protein serine/threonine phosphatases and the protein tyrosine phosphatases (PTPs; protein-tyrosine-phosphate phosphohydrolase, EC 3.13.48. Hunter, 1989. The PTPs are a family of proteins that have been classified into two subgroups. The first subgroup is made up of the low molecular weight, intracellular enzymes that contain a single conserved catalytic phosphatase domain. All known intracellular type PTPs contain a single conserved catalytic phosphatase domain. Examples of the first group of PTPs include (1) placental PTP 1B (Charbonneau *et al.*, *Proc. Natl. Acad. Sci. USA 86*, 5252-56, 1989; Chernoff *et al.*, *Proc. Natl. Acad. Sci. USA 87*, 2735-89, 1990, (2) T-cell PTP (Cool *et al.*, *Proc. Natl. Acad. Sci. USA 86*, 5257-61, 1989, (3) rat brain PTP (Guan *et al.*, *Proc. Natl. Acad. Sci. USA 87*, 1501-02, 1990,

(4) neuronal phosphatase (STEP) (Lombroso *et al.*, *Proc. Natl. Acad. Sci. USA 88*, 7242-46, 1991, and (5) cytoplasmic phosphatases that contain a region of homology to cytoskeletal proteins (Gu *et al.*, *Proc. Natl. Acad. Sci. USA 88*, 5867-871, 1991; Yang *et al.*, *Proc. Natl. Acad. Sci. USA 88*, 5949-53, 1991.

The second subgroup is made up of the high molecular weight, receptor-linked PTPs, termed RPTPs. RPTPs consist of (a) an intracellular catalytic region, (b) a single transmembrane segment, and (c) a putative ligand-binding extracellular domain. The structures and sizes of the putative ligand-binding extracellular "receptor" domains of RPTPs are quite divergent. In contrast, the intracellular catalytic regions of RPTPs are highly homologous. All RPTPs have two tandemly duplicated catalytic phosphatase homology domains, with the prominent exception of an RPTP termed HPTPβ, which has only one catalytic phosphatase domain. Tsai *et al.*, *J. Biol. Chem. 266*, 0534-43, 1991.

Since the initial purification, sequencing, and cloning of a PTP, additional potential PTPs have been identified at a rapid pace. The number of different PTPs that have been identified is increasing steadily, leading to speculations that this family may be as large as the PTK family. Hunter, 1989.

Conserved amino acid sequences designated "consensus sequences" have been identified in the catalytic domains of known PTPs. Krueger *et al.*, *EMBO J. 9*, 3241-52, 1990; Yi *et al.*, *Mol. Cell. Biol. 12*, 836-46, 1992. Yi *et al.* aligned the catalytic phosphatase domain sequences of the following PTPs: LCA, PTP1B, TCPTP, LAR, DLAR, and HPTPα, HPTPβ, and HPTPγ. This alignment includes the following consensus sequences (Yi *et al.*, *supra*):

1. K C X X Y W P
2. H C S X G X G R X G.

Krueger et al., aligned the catalytic phosphatase domain sequences of PTP1B, TCPTP, LAR, LCA, HPTPα, β, δ, ε and ζ, and DLAR and DPTP.

This alignment includes the following consensus sequences (Krueger *et al.*, *supra*):

1. K C X X Y W P

2. H C S X G X G R X G.

It is becoming clear that dephosphorylation of tyrosine residues can by itself function as an important regulatory mechanism. Dephosphorylation of a C-terminal tyrosine residue has been shown to activate tyrosine kinase activity in the src family of tyrosine kinases. Hunter, *Cell 49*, 1-4, 1987. Tyrosine dephosphorylation has been suggested to be an obligatory step in the mitotic activation of the maturation-promoting factor (MPF) kinase. Morla *et al.*, *Cell 58*, 193-203, 1989. These observations point out the need in the art for understanding the mechanisms that regulate tyrosine phosphatase activity.

Because of the importance of these enzymes in a variety of cellular processes, there is a continuing need in the art to identify additional tyrosine phosphatases, which can be regulated to provide therapeutic effects.

BRIEF SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human receptor tyrosine phosphatase. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a receptor tyrosine phosphatase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 87% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a receptor tyrosine phosphatase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 87% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

Binding between the test compound and the receptor tyrosine phosphatase polypeptide is detected. A test compound which binds to the receptor tyrosine phosphatase polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the receptor tyrosine phosphatase.

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a receptor tyrosine phosphatase polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the

receptor tyrosine phosphatase through interacting with the receptor tyrosine phosphatase mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a receptor tyrosine phosphatase polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 87% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

A receptor tyrosine phosphatase activity of the polypeptide is detected. A test compound which increases receptor tyrosine phosphatase activity of the polypeptide relative to receptor tyrosine phosphatase activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases receptor tyrosine phosphatase activity of the polypeptide relative to receptor tyrosine phosphatase activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a receptor tyrosine phosphatase product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

Binding of the test compound to the receptor tyrosine phosphatase product is detected. A test compound which binds to the receptor tyrosine phosphatase product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a receptor tyrosine phosphatase polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

Receptor tyrosine phosphatase activity in the cell is thereby decreased.

The invention thus provides a human receptor tyrosine phosphatase that can be used to identify test compounds that may act, for example, as activators or inhibitors at the enzyme's active site. Human receptor tyrosine phosphatase and fragments thereof also are useful in raising specific antibodies that can block the enzyme and effectively reduce its activity.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 shows the DNA-sequence encoding a receptor tyrosine phosphatase Polypeptide (SEQ ID NO:1).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO:2).

Fig. 3 shows the amino acid sequence of the protein identified by trembl|AF063249|AF063249_1 (SEQ ID NO:3).

Fig. 4 shows the amino acid sequence of a receptor tyrosine phosphatase Polypeptide (SEQ ID NO:4).

Fig. 5 shows the amino acid sequence of a receptor tyrosine phosphatase Polypeptide (SEQ ID NO:5).

Fig. 6 shows the DNA-sequence encoding a receptor tyrosine phosphatase Polypeptide (SEQ ID NO:6).

Fig. 7 shows the DNA-sequence encoding a receptor tyrosine phosphatase Polypeptide (SEQ ID NO:7).

Fig. 8 shows the DNA-sequence encoding a receptor tyrosine phosphatase Polypeptide (SEQ ID NO:8).

Fig. 9 shows the BLASTP - alignment of 492_Protein (SEQ ID NO:2) against trembl|AF063249|AF063249_1 (SEQ ID NO:3). This hit is scoring at: 0.0 (expectation value); Alignment length (overlap): 2301 ; Identities : 86 % ; Scoring matrix : BLOSUM62 (used to infer consensus pattern) ; Database searched : nrdb_1_.

Fig. 10 shows the BLASTP - alignment of 492_Protein (SEQ ID NO:2) against pdb|1LAR|1LAR-A larfragment: phosphatase. This hit is scoring at: 3e-60 (expectation value); Alignment length (overlap): 308; Identities: 40 % ; Scoring matrix : BLOSUM62 (used to infer consensus pattern) ; Database searched : nrdb_1_.

Fig. 11 shows the BLASTP – alternative alignment of 492_Protein (SEQ ID NO:2) against pdb|1LAR|1LAR-A lar fragment: phosphatase. This hit is scoring at: 3e-44 (expectation value); Alignment length (overlap): 245; Identities: 40 %; Scoring matrix: BLOSUM62 (used to infer consensus pattern); Database searched: nrdb_1_.

Fig. 12 shows the HMMPFAM - alignment of 492_Protein (SEQ ID NO:2) against pfam|hmm|Y_phosphatase Protein-tyrosine phosphatase. This hit is scoring at: 374.0; Scoring matrix: BLOSUM62 (used to infer consensus pattern).

Fig. 13 shows the ATP_GTP_A region, transmembrane segments, and fibronectin type III domain region prediction. ATP_GTP_A region from residue 1621 to 1629. Source: [prosite]; Tyrosine specific protein phosphatases signature; -Consensus pattern: [LIVMF]-H-C-x(2)-G-x(3)-[STC]-[STAGP]-x-[LIVMFY] [C is the active site residue]; ATP/GTP-binding site motif A (P-loop); -Consensus pattern: [AG]-x(4)-G-K-[ST]; High scoring transmembrane segments: From 1906 to 1928: length= 23, score=69.00 ** 1906 IILSVTLCIL SIILLGTAIF AFAL: 5(21.7%); A: 3(13.0%); I: 6(26.1%); Fibronectin type III domain region.

Fig. 14 shows the Genewise output: use rat "glomerular mesangial cell receptor protein-tyrosine phosphatase precursor" (AF063249) (SEQ ID NO:3) as a template on human genomic DNA (NT_019546; chr12:8890000188900001-32923-88900001269821); Score 3940.84 bits over entire alignment; Scores as bits over a synchronous coding model.

Fig. 15 shows the Genewise output: use partial protein (492_BAC_p, below) of rat "glomerular mesangial cell receptor protein-tyrosine phosphatase precursor" (AF063249) (SEQ ID NO:3) as a template on human 12q BAC RP11-288D9 (gi|15011617|gb|AC083812.24).

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide from the group consisting of:

a) a polynucleotide encoding a receptor tyrosine phosphatase polypeptide comprising an amino acid sequence selected from the group consisting of:
amino acid sequences which are at least about 87% identical to the amino acid sequence shown in SEQ ID NO: 2; and
the amino acid sequence shown in SEQ ID NO: 2.

b) a polynucleotide comprising the sequence of SEQ ID NO: 1;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a receptor tyrosine phosphatase polypeptide;

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a receptor tyrosine phosphatase polypeptide; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a receptor tyrosine phosphatase polypeptide.

Furthermore, it has been discovered by the present applicant that a novel receptor tyrosine phosphatase, particularly a human receptor tyrosine phosphatase, can be used in therapeutic methods to treat diabetes, a CNS disorder, COPD or a cardiovascular disorder. Human receptor tyrosine phosphatase comprises the amino acid sequence shown in SEQ ID NO:2. A coding sequence for human receptor tyrosine phosphatase is shown in SEQ ID NO:1. This sequence is located on chromosome 12. Related ESTs (AF169351; AR073855; AR073854; AR073853; AR073852; AR073851; AR073850) are expressed in colon and K562 CML cell line.

Human receptor tyrosine phosphatase is 86% identical over 2301 amino acids to glomerular mesangial cell receptor protein-tyrosine phosphatase precursor (FIG. 1). Human receptor tyrosine phosphatase of the invention is expected to be useful for the same purposes as previously identified receptor tyrosine phosphatase enzymes.

Human receptor tyrosine phosphatase is believed to be useful in therapeutic methods to treat disorders such as diabetes, obesity, a CNS disorder, COPD, and cardiovascular disorders. Human receptor tyrosine phosphatase also can be used to screen for human receptor tyrosine phosphatase activators and inhibitors.

Polypeptides

Human receptor tyrosine phosphatase polypeptides according to the invention comprise at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, 1700, 1725, 1750, 1775, 1800, 1825, 1850, 1875, 1900, 1925, 1950, 1975, 2000, 2025, 2050, 2075, 2100, 2125, 2150, 2175, 2200, 2225, 2250, 2275 or 2299 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO:2 or a biologically active variant thereof, as defined below. A receptor tyrosine phosphatase polypeptide of the invention therefore can be a portion of a receptor tyrosine phosphatase protein, a full-length receptor tyrosine phosphatase protein, or a fusion protein comprising all or a portion of a receptor tyrosine phosphatase protein.

Biologically active variants

Human receptor tyrosine phosphatase polypeptide variants that are biologically active, *e.g.*, retain an enzymatic activity, also are receptor tyrosine phosphatase polypeptides. Preferably, naturally or non-naturally occurring receptor tyrosine

phosphatase polypeptide variants have amino acid sequences which are at least about 87, preferably about 90, 96, 96, 98, or 99% identical to the amino acid sequence shown in SEQ ID NO:2 or a fragment thereof. Percent identity between a putative receptor tyrosine phosphatase polypeptide variant and an amino acid sequence of SEQ ID NO:2 is determined using the Blast2 alignment program (Blosum62, Expect 10, standard genetic codes)..

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a human receptor tyrosine phosphatase polypeptide can be found using computer programs well known in the art, such as DNASTAR software.

The invention additionally, encompasses receptor tyrosine phosphatase polypeptides that are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications can be carried out by known techniques including, but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, etc.

Additional post-translational modifications encompassed by the invention include, for example, *e.g.*, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The receptor tyrosine phosphatase polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

The invention also provides chemically modified derivatives of receptor tyrosine phosphatase polypeptides that may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization can be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol, and the like. The polypeptides can be modified at random or predetermined positions within the molecule and can include one, two, three, or more attached chemical moieties.

Whether an amino acid change or a polypeptide modification results in a biologically active receptor tyrosine phosphatase polypeptide can readily be determined by assaying for enzymatic activity, as described for example, in Calvert-Evers & Hammond, *Cell Biol. Int. 24*, 559-68, 2000.

Fusion proteins

Fusion proteins are useful for generating antibodies against receptor tyrosine phosphatase polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins that interact with portions of a human receptor tyrosine phosphatase polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the

yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A human receptor tyrosine phosphatase polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, 1700, 1725, 1750, 1775, 1800, 1825, 1850, 1875, 1900, 1925, 1950, 1975, 2000, 2025, 2050, 2075, 2100, 2125, 2150, 2175, 2200, 2225, 2250, 2275 or 2299 contiguous amino acids of SEQ ID NO:2 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length receptor tyrosine phosphatase protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the receptor tyrosine phosphatase polypeptide-encoding sequence and the heterologous protein sequence, so that the receptor tyrosine phosphatase polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO:1 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of species homologs

Species homologs of human receptor tyrosine phosphatase polypeptide can be obtained using receptor tyrosine phosphatase polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of receptor tyrosine phosphatase polypeptide, and expressing the cDNAs as is known in the art.

Polynucleotides

A human receptor tyrosine phosphatase polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a receptor tyrosine phosphatase polypeptide. A coding sequence for human receptor tyrosine phosphatase is shown in SEQ ID NO:1.

Degenerate nucleotide sequences encoding human receptor tyrosine phosphatase polypeptides, as well as homologous nucleotide sequences which are at least about

50, 55, 60, 65, 70, preferably about 75, 90, 96, 98, or 99% identical to the nucleotide sequence shown in SEQ ID NO:1 or its complement also are receptor tyrosine phosphatase polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of receptor tyrosine phosphatase polynucleotides that encode biologically active receptor tyrosine phosphatase polypeptides also are receptor tyrosine phosphatase polynucleotides. Polynucleotide fragments comprising at least 8, 9, 10, 11, 12, 15, 20, or 25 contiguous nucleotides of SEQ ID NO:1 or its complement also are receptor tyrosine phosphatase polynucleotides. These fragments can be used, for example, as hybridization probes or as antisense oligonucleotides.

Identification of polynucleotide variants and homologs

Variants and homologs of the receptor tyrosine phosphatase polynucleotides described above also are receptor tyrosine phosphatase polynucleotides. Typically, homologous receptor tyrosine phosphatase polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known receptor tyrosine phosphatase polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 °C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the receptor tyrosine phosphatase polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening

cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of receptor tyrosine phosphatase polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al., J. Mol. Biol. 81*, 123 (1973). Variants of human receptor tyrosine phosphatase polynucleotides or receptor tyrosine phosphatase polynucleotides of other species can therefore be identified by hybridizing a putative homologous receptor tyrosine phosphatase polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to receptor tyrosine phosphatase polynucleotides or their complements following stringent hybridization and/or wash conditions also are receptor tyrosine phosphatase polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20 °C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a receptor tyrosine phosphatase polynucleotide having a nucleotide sequence shown in SEQ ID NO:1 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

T_m = 81.5 °C - 16.6($\log_{10}$[Na^+]) + 0.41(%G + C) - 0.63(%formamide) - 600/*l*), where *l* = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65 °C, or 50% formamide, 4X SSC at 42 °C, or 0.5X SSC, 0.1% SDS at 65 °C. Highly stringent wash conditions include, for example, 0.2X SSC at 65 °C.

Preparation of polynucleotides

A human receptor tyrosine phosphatase polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated receptor tyrosine phosphatase polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments, which comprise receptor tyrosine phosphatase nucleotide sequences. Isolated polynucleotides are in preparations that are free or at least 70, 80, or 90% free of other molecules.

Human receptor tyrosine phosphatase cDNA molecules can be made with standard molecular biology techniques, using receptor tyrosine phosphatase mRNA as a template. Human receptor tyrosine phosphatase cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize receptor tyrosine phosphatase polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a human receptor tyrosine phosphatase polypeptide having, for example, an amino acid sequence shown in SEQ ID NO:2 or a biologically active variant thereof.

Extending polynucleotides

Both methods are well known in the art. Briefly, filters with bacterial colonies containing the library in pBluescript or bacterial lawns containing lambda plaques are denatured, and the DNA is fixed to the filters. The filters are hybridized with the labeled probe using hybridization conditions described by Davis *et al.*, 1986. The partial sequences, cloned into lambda or pBluescript, can be used as positive controls to assess background binding and to adjust the hybridization and washing stringencies necessary for accurate clone identification. The resulting autoradiograms are compared to duplicate plates of colonies or plaques; each exposed spot corresponds to a positive colony or plaque. The colonies or plaques are selected, expanded and the DNA is isolated from the colonies for further analysis and sequencing.

Positive cDNA clones are analyzed to determine the amount of additional sequence they contain using PCR with one primer from the partial sequence and the other primer from the vector. Clones with a larger vector-insert PCR product than the original partial sequence are analyzed by restriction digestion and DNA sequencing to determine whether they contain an insert of the same size or similar as the mRNA size determined from Northern blot Analysis.

Once one or more overlapping cDNA clones are identified, the complete sequence of the clones can be determined, for example after exonuclease III digestion (McCombie *et al.*, *Methods 3*, 33-40, 1991). A series of deletion clones are generated, each of which is sequenced. The resulting overlapping sequences are

assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a highly accurate final sequence.

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al., Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template. Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al., PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al., Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) that are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA that might be present in limited amounts in a particular sample.

Obtaining Polynucleotides

Receptor tyrosine phosphatase polypeptides can be obtained, for example, by purification from human cells, by expression of receptor tyrosine phosphatase polynucleotides, or by direct chemical synthesis.

Protein purification

Receptor tyrosine phosphatase polypeptides can be purified from any human cell which expresses the receptor, including host cells which have been transfected with receptor tyrosine phosphatase polynucleotides. A purified receptor tyrosine phosphatase polypeptide is separated from other compounds that normally associate with the receptor tyrosine phosphatase polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis.

Receptor tyrosine phosphatase polypeptide can be conveniently isolated as a complex with its associated G protein, as described in the specific examples, below. A preparation of purified receptor tyrosine phosphatase polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of polynucleotides

To express a human receptor tyrosine phosphatase polynucleotide, the polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding receptor tyrosine phosphatase polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a human receptor tyrosine phosphatase polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life

Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a human receptor tyrosine phosphatase polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and yeast expression systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the receptor tyrosine phosphatase polypeptide. For example, when a large quantity of a human receptor tyrosine phosphatase polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the receptor tyrosine phosphatase polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516-544, 1987.

Plant and insect expression systems

If plant expression vectors are used, the expression of sequences encoding receptor tyrosine phosphatase polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 1671-1680, 1984; Broglie *et al.*, *Science 224*, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in McGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a human receptor tyrosine phosphatase polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding receptor tyrosine phosphatase polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of receptor tyrosine phosphatase polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which receptor tyrosine phosphatase polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

Mammalian expression systems

A number of viral-based expression systems can be used to express receptor tyrosine phosphatase polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding receptor tyrosine phosphatase polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a human receptor tyrosine phosphatase polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci. 81*, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding receptor tyrosine phosphatase polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a human receptor tyrosine phosphatase polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell

system which is used (see Scharf *et al., Results Probl. Cell Differ. 20*, 125-162, 1994).

Host cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed receptor tyrosine phosphatase polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells that have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express receptor tyrosine phosphatase polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced receptor tyrosine phosphatase sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 817-23, 1980) genes that can be employed in *tk⁻* or *aprt⁻* cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol. 55*, 121-131, 1995).

Detecting expression

Although the presence of marker gene expression suggests that the receptor tyrosine phosphatase polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a human receptor tyrosine phosphatase polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode an receptor tyrosine phosphatase polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a receptor tyrosine phosphatase polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the receptor tyrosine phosphatase polynucleotide.

Alternatively, host cells which contain a human receptor tyrosine phosphatase polynucleotide and which express a human receptor tyrosine phosphatase polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques that include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding an receptor tyrosine phosphatase polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a human receptor tyrosine phosphatase polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a receptor tyrosine phosphatase polypeptide to detect transformants which contain an receptor tyrosine phosphatase polynucleotide.

A variety of protocols for detecting and measuring the expression of a human receptor tyrosine phosphatase polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a human receptor tyrosine phosphatase polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding receptor tyrosine phosphatase polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled

nucleotide. Alternatively, sequences encoding a human receptor tyrosine phosphatase polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and purification of polypeptides

Host cells transformed with nucleotide sequences encoding a human receptor tyrosine phosphatase polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode receptor tyrosine phosphatase polypeptides can be designed to contain signal sequences which direct secretion of soluble receptor tyrosine phosphatase polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound receptor tyrosine phosphatase polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a human receptor tyrosine phosphatase polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system

(Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the receptor tyrosine phosphatase polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a human receptor tyrosine phosphatase polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the receptor tyrosine phosphatase polypeptide from the fusion protein. Vectors that contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol. 12*, 441-453, 1993.

Chemical synthesis

Sequences encoding a human receptor tyrosine phosphatase polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a human receptor tyrosine phosphatase polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al.*, *Science 269*, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of receptor tyrosine phosphatase polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The

composition of a synthetic receptor tyrosine phosphatase polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the receptor tyrosine phosphatase polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

As will be understood by those of skill in the art, it may be advantageous to produce receptor tyrosine phosphatase polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter receptor tyrosine phosphatase polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a human receptor tyrosine phosphatase polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a human receptor

tyrosine phosphatase polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a human receptor tyrosine phosphatase polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody that specifically binds to the immunogen.

Typically, an antibody which specifically binds to a human receptor tyrosine phosphatase polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to receptor tyrosine phosphatase polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a human receptor tyrosine phosphatase polypeptide from solution.

Human receptor tyrosine phosphatase polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a human receptor tyrosine phosphatase polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants

used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies which specifically bind to a human receptor tyrosine phosphatase polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature 256*, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608, 1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies that specifically bind to a human receptor tyrosine phosphatase polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies that

specifically bind to receptor tyrosine phosphatase polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91).

Antibodies which specifically bind to receptor tyrosine phosphatase polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and

which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a human receptor tyrosine phosphatase polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of receptor tyrosine phosphatase gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of receptor tyrosine phosphatase gene expression can be obtained by designing antisense oligonucleotides that will form duplexes to the control, 5', or regulatory regions of the receptor tyrosine phosphatase gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a human receptor tyrosine phosphatase polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to an receptor tyrosine phosphatase polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent receptor tyrosine phosphatase nucleotides, can provide sufficient targeting specificity for receptor tyrosine phosphatase mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular receptor tyrosine phosphatase polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a human receptor tyrosine phosphatase polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For

example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a human receptor tyrosine phosphatase polynucleotide can be used to generate ribozymes that will specifically bind to mRNA transcribed from the receptor tyrosine phosphatase polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to

the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a human receptor tyrosine phosphatase RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate receptor tyrosine phosphatase RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease receptor tyrosine phosphatase expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors that induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of

regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Differentially expressed genes

Described herein are methods for the identification of genes whose products interact with human receptor tyrosine phosphatase. Such genes may represent genes that are differentially expressed in disorders including, but not limited to, diabetes, CNS disorders, COPD, and cardiovascular disorders. Further, such genes may represent genes that are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human receptor tyrosine phosphatase gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique that does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed

using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples that represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human receptor tyrosine phosphatase. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human receptor tyrosine phosphatase. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human receptor tyrosine phosphatase gene or gene product are up-regulated or down-regulated.

Screening methods

The invention provides assays for screening test compounds that bind to or modulate the activity of a human receptor tyrosine phosphatase polypeptide or a human receptor tyrosine phosphatase polynucleotide. A test compound preferably binds to a human receptor tyrosine phosphatase polypeptide or polynucleotide. More preferably, a test compound decreases or increases enzymatic activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution (*see, e.g.*, Houghten, *BioTechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High through-put screening

Test compounds can be screened for the ability to bind to receptor tyrosine phosphatase polypeptides . or polynucleotides or to affect receptor tyrosine phosphatase activity or receptor tyrosine phosphatase gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19,* 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding assays

For binding assays, the test compound is preferably a small molecule that binds to and occupies, for example, the active site of the receptor tyrosine phosphatase polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the receptor tyrosine phosphatase polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound that is bound to the receptor tyrosine phosphatase polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a human receptor tyrosine phosphatase polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with

a human receptor tyrosine phosphatase polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a human receptor tyrosine phosphatase polypeptide (McConnell *et al.*, *Science 257*, 1906-1912, 1992).

Determining the ability of a test compound to bind to a human receptor tyrosine phosphatase polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a human receptor tyrosine phosphatase polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 12046-12054, 1993; Bartel *et al.*, *BioTechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent W094/10300), to identify other proteins which bind to or interact with the receptor tyrosine phosphatase polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a human receptor tyrosine phosphatase polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation

domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein that interacts with the receptor tyrosine phosphatase polypeptide.

It may be desirable to immobilize either the receptor tyrosine phosphatase polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the receptor tyrosine phosphatase polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a human receptor tyrosine phosphatase polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the receptor tyrosine phosphatase polypeptide is a fusion protein comprising a domain that allows the receptor tyrosine phosphatase polypeptide to be bound to a solid support.. For example, glutathione-S-transferase fusion proteins can

be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed receptor tyrosine phosphatase polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a human receptor tyrosine phosphatase polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated receptor tyrosine phosphatase polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N- hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a receptor tyrosine phosphatase polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the receptor tyrosine phosphatase polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the receptor tyrosine phosphatase polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the receptor tyrosine phosphatase polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a human receptor tyrosine phosphatase polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a receptor tyrosine phosphatase polypeptide or polynucleotide can be used in a cell-based assay system. A receptor tyrosine phosphatase polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a receptor tyrosine phosphatase polypeptide or polynucleotide is determined as described above.

Enzymatic activity

Test compounds can be tested for the ability to increase or decrease the enzymatic activity of a human receptor tyrosine phosphatase polypeptide. Enzymatic activity can be measured, for example, as described in Calvert-Evers & Hammond, *Cell Biol. Int. 24*, 559-68, 2000.

Enzyme assays can be carried out after contacting either a purified receptor tyrosine phosphatase polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test compound that decreases enzymatic activity of a human receptor tyrosine phosphatase polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing receptor tyrosine phosphatase activity. A test compound which increases enzymatic activity of a human receptor tyrosine phosphatase polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing human receptor tyrosine phosphatase activity.

Gene expression

In another embodiment, test compounds that increase or decrease receptor tyrosine phosphatase gene expression are identified. A receptor tyrosine phosphatase polynucleotide is contacted with a test compound, and the expression of an RNA or

polypeptide product of the receptor tyrosine phosphatase polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of receptor tyrosine phosphatase mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a human receptor tyrosine phosphatase polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a human receptor tyrosine phosphatase polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell that expresses a human receptor tyrosine phosphatase polynucleotide can be used in a cell-based assay system. The receptor tyrosine phosphatase polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical compositions

The invention also provides pharmaceutical compositions that can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a human receptor tyrosine phosphatase polypeptide, receptor tyrosine phosphatase polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a receptor tyrosine phosphatase polypeptide, or mimetics, activators, or inhibitors of a human receptor tyrosine phosphatase polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to

obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides,

or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactio, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic indications and methods

Human receptor tyrosine phosphatase can be regulated to treat diabetes, CNS disorders, COPD, and cardiovascular disorders.

Diabetes mellitus is a common metabolic disorder characterized by an abnormal elevation in blood glucose, alterations in lipids and abnormalities (complications) in the cardiovascular system, eye, kidney and nervous system. Diabetes is divided into two separate diseases: type 1 diabetes (juvenile onset), which results from a loss of cells which make and secrete insulin, and type 2 diabetes (adult onset), which is caused by a defect in insulin secretion and a defect in insulin action.

Type 1 diabetes is initiated by an autoimmune reaction that attacks the insulin secreting cells (beta cells) in the pancreatic islets. Agents that prevent this reaction from occurring or that stop the reaction before destruction of the beta cells has been accomplished are potential therapies for this disease. Other agents that induce beta cell proliferation and regeneration also are potential therapies.

Type II diabetes is the most common of the two diabetic conditions (6% of the population). The defect in insulin secretion is an important cause of the diabetic condition and results from an inability of the beta cell to properly detect and respond to rises in blood glucose levels with insulin release. Therapies that increase the response by the beta cell to glucose would offer an important new treatment for this disease.

The defect in insulin action in Type II diabetic subjects is another target for therapeutic intervention. Agents that increase the activity of the insulin receptor in muscle, liver, and fat will cause a decrease in blood glucose and a normalization of plasma lipids. The receptor activity can be increased by agents that directly stimulate the receptor or that increase the intracellular signals from the receptor. Other therapies can directly activate the cellular end process, *i.e.* glucose transport or various enzyme systems, to generate an insulin-like effect and therefore a produce beneficial outcome. Because overweight subjects have a greater susceptibility to Type II diabetes, any agent that reduces body weight is a possible therapy.

Both Type I and Type diabetes can be treated with agents that mimic insulin action or that treat diabetic complications by reducing blood glucose levels. Likewise, agents that reduces new blood vessel growth can be used to treat the eye complications that develop in both diseases.

Central and peripheral nervous system disorders also can be treated, such as primary and secondary disorders after brain injury, disorders of mood, anxiety disorders, disorders of thought and volition, disorders of sleep and wakefulness, diseases of the motor unit, such as neurogenic and myopathic disorders, neurodegenerative disorders such as Alzheimer's and Parkinson's disease, and processes of peripheral and chronic pain.

Pain that is associated with CNS disorders also can be treated by regulating the activity of human receptor tyrosine phosphatase. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (*e.g.*, infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgiaradioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (*e.g.*, diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

Central and peripheral nervous system disorders also can be treated, such as primary and secondary disorders after brain injury, disorders of mood, anxiety disorders, disorders of thought and volition, disorders of sleep and wakefulness, diseases of the motor unit, such as neurogenic and myopathic disorders, neurodegenerative disorders such as Alzheimer's and Parkinson's disease, and processes of peripheral and chronic pain.

Pain that is associated with CNS disorders also can be treated by regulating the activity of human receptor tyrosine phosphatase. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (*e.g.*, infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgiaradioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (*e.g.*, diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

Cardiovascular diseases include the following disorders of the heart and the vascular system: congestive heart failure, myocardial infarction, ischemic diseases of the heart, all kinds of atrial and ventricular arrhythmias, hypertensive vascular diseases, and peripheral vascular diseases.

Heart failure is defined as a pathophysiologic state in which an abnormality of cardiac function is responsible for the failure of the heart to pump blood at a rate commensurate with the requirement of the metabolizing tissue. It includes all forms of pumping failure, such as high-output and low-output, acute and chronic, right-sided or left-sided, systolic or diastolic, independent of the underlying cause.

Myocardial infarction (MI) is generally caused by an abrupt decrease in coronary blood flow that follows a thrombotic occlusion of a coronary artery previously narrowed by arteriosclerosis. MI prophylaxis (primary and secondary prevention) is included, as well as the acute treatment of MI and the prevention of complications.

Ischemic diseases are conditions in which the coronary flow is restricted resulting in a perfusion which inadequate to meet the myocardial requirement for oxygen. This group of diseases includes stable angina, unstable angina, and asymptomatic ischemia.

Arrhythmias include all forms of atrial and ventricular tachyarrhythmias (atrial tachycardia, atrial flutter, atrial fibrillation, atrio-ventricular reentrant tachycardia, preexcitation syndrome, ventricular tachycardia, ventricular flutter, and ventricular fibrillation), as well as bradycardic forms of arrhythmias.

Vascular diseases include primary as well as all kinds of secondary arterial hypertension (renal, endocrine, neurogenic, others). The disclosed gene and its product may be used as drug targets for the treatment of hypertension as well as for the prevention of all complications.

Peripheral vascular diseases are defined as vascular diseases in which arterial and/or venous flow is reduced resulting in an imbalance between blood supply and tissue oxygen demand. It includes chronic peripheral arterial occlusive disease (PAOD), acute arterial thrombosis and embolism, inflammatory vascular disorders, Raynaud's phenomenon, and venous disorders.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an anti-sense nucleic acid molecule, a specific antibody, ribozyme, or a human receptor tyrosine phosphatase polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects receptor tyrosine phosphatase activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce receptor tyrosine phosphatase activity. The reagent preferably binds to an expression product of a human receptor tyrosine phosphatase gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells that have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods that are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988);

Wu *et al.*, *J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem. 266*, 338-42 (1991).

Determination of a therapeutically effective dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases enzymatic activity relative to the enzymatic activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions that exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to

provide sufficient levels of the active ingredient or to maintain the desired effect. Factors that can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides that express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a human receptor tyrosine phosphatase gene or the activity of a receptor tyrosine phosphatase polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a human receptor tyrosine phosphatase gene or the activity of a human receptor tyrosine phosphatase polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to receptor tyrosine phosphatase-specific mRNA, quantitative RT-PCR, immunologic detection of a human receptor tyrosine phosphatase polypeptide, or measurement of enzymatic activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic methods

Human receptor tyrosine phosphatase also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences that encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding receptor tyrosine phosphatase in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al.*, *Science 230*, 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA 85*, 4397-4401, 1985). Thus, the detection of a specific DNA sequence can be performed

by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of receptor tyrosine phosphatase also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of receptor tyrosine phosphatase activity

The polynucleotide of SEQ ID NO: 1 is inserted into the expression vector pCEV4 and the expression vector pCEV4-receptor tyrosine phosphatase polypeptide obtained is transfected into human embryonic kidney 293 cells. To prepare membrane fractions, the cells are sonicated in a hypotonic lysis buffer (25 mM Tris-HCl, pH 7.5/25 mM sucrose/0.1 mM EDTA/5 mM MgC12/5 mM DTT/1 mM phenylmethylsulfonyl fluoride/0.5 µg/ml leupeptin/1 µg/ml aprotinin), nuclei are removed by low-speed centrifugation, and membrane fractions are obtained by centrifugation at 100,000 x g for 60 min at 4°C. The resulting pellets are suspended by sonication in lysis buffer, brought to a concentration of 2 mg/ml, and used to measure receptor tyrosine phosphatase activity.. A nonradioactive protein-tyrosine phosphatase assay kit (Roce Molecular Biochemicals) is used according to the

manufacturer's instructions. Signals on Western blots are detected by chemiluminescence (ECL, Amersham Pharmacia Biotech. It is shown that the polypeptide of SEQ ID NO: 2 has a receptor tyrosine phosphatase activity.

EXAMPLE 2

Expression of recombinant human receptor tyrosine phosphatase

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human receptor tyrosine phosphatase polypeptides in yeast. The receptor tyrosine phosphatase-encoding DNA sequence is derived from SEQ ID NO:1. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human receptor tyrosine phosphatase polypeptide is obtained.

EXAMPLE 3

Identification of test compounds that bind to receptor tyrosine phosphatase polypeptides

Purified receptor tyrosine phosphatase polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well micro-titer plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Human receptor tyrosine phosphatase polypeptides comprise the amino acid sequence shown in SEQ ID NO:2. The test compounds comprise a fluorescent tag.. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a human receptor tyrosine phosphatase polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound that increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a human receptor tyrosine phosphatase polypeptide.

EXAMPLE 4

Identification of a test compound which decreases receptor tyrosine phosphatase gene expression

A test compound is administered to a culture of human cells transfected with a receptor tyrosine phosphatase expression construct and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control. RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 μg total RNA and hybridized with a ^{32}P-labeled receptor tyrosine phosphatase-specific probe at 65 ° C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous

nucleotides selected from the complement of SEQ ID NO:1. A test compound that decreases the receptor tyrosine phosphatase-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of receptor tyrosine phosphatase gene expression.

EXAMPLE 5

Identification of a test compound which decreases receptor tyrosine phosphatase activity

A test compound is administered to a culture of human cells transfected with a receptor tyrosine phosphatase expression construct and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control. Enzymatic activity is measured using the method of Calvert-Evers & Hammond, *Cell Biol. Int. 24*, 559-68, 2000.

A test compound which decreases the enzymatic activity of the receptor tyrosine phosphatase relative to the enzymatic activity in the absence of the test compound is identified as an inhibitor of receptor tyrosine phosphatase activity.

EXAMPLE 6

Tissue-specific expression of receptor tyrosine phosphatase

The qualitative expression pattern of receptor tyrosine phosphatase in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR).

Quantitative expression profiling.

To demonstrate that receptor tyrosine phosphatase is involved in the disease process of diabetes, the following whole body panel is screened to show predominant or relatively high expression: subcutaneous and mesenteric adipose tissue, adrenal

gland, bone marrow, brain, colon, fetal brain, heart, hypothalamus, kidney, liver, lung, mammary gland, pancreas, placenta, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, thyroid, trachea, and uterus. Human islet cells and an islet cell library also are tested. As a final step, the expression of receptor tyrosine phosphatase in cells derived from normal individuals with the expression of cells derived from diabetic individuals is compared.

To demonstrate that receptor tyrosine phosphatase is involved in CNS disorders, the following tissues are screened: fetal and adult brain, muscle, heart, lung, kidney, liver, thymus, testis, colon, placenta, trachea, pancreas, kidney, gastric mucosa, colon, liver, cerebellum, skin, cortex (Alzheimer's and normal), hypothalamus, cortex, amygdala, cerebellum, hippocampus, choroid, plexus, thalamus, and spinal cord.

To demonstrate that receptor tyrosine phosphatase is involved in the disease process of COPD, the initial expression panel consists of RNA samples from respiratory tissues and inflammatory cells relevant to COPD: lung (adult and fetal), trachea, freshly isolated alveolar type II cells, cultured human bronchial epithelial cells, cultured small airway epithelial cells, cultured bronchial sooth muscle cells, cultured H441 cells (Clara-like), freshly isolated neutrophils and monocytes, and cultured monocytes (macrophage-like). Body map profiling also is carried out, using total RNA panels purchased from Clontech. The tissues are adrenal gland, bone marrow, brain, colon, heart, kidney, liver, lung, mammary gland, pancreas, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, trachea, thyroid, and uterus.

Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in Higuchi *et al.*, *BioTechnology 10*, 413-17, 1992; and Higuchi *et al.*, *BioTechnology 11*, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al., Proc. Natl. Acad. Sci. U.S.A. 88*, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al., Genome Res. 6*, 986-94, 1996, and Gibson *et al., Genome Res. 6*, 995-1001, 1996).

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

RNA extraction and cDNA preparation. Total RNA from the tissues listed above are used for expression quantification. RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10mM $MgCl_2$; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:iso-amyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M sodium acetate, pH5.2, and 2 volumes of ethanol.

Fifty μg of each RNA from the autoptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophoto-metric quantification, each sample is reverse transcribed with the TaqMan Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manu-facturer's protocol. The final concentration of RNA in the reaction mix is 200ng/μL. Reverse transcription is carried out with 2.5μM of random hexamer primers.

TaqMan quantitative analysis. Specific primers and probe are designed according to the recommendations of PE Applied Biosystems; the probe can be labeled at the 5' end FAM (6-carboxy-fluorescein) and at the 3' end with TAMRA (6-carboxy-tetra-methyl-rhodamine). Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 μl.

Each of the following steps are carried out once: pre PCR, 2 minutes at 50° C, and 10 minutes at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

EXAMPLE 7

Diabetes: In vivo testing of compounds/target validation

Glucose Production

Over-production of glucose by the liver, due to an enhanced rate of gluconeogenesis, is the major cause of fasting hyperglycemia in diabetes. Overnight fasted normal rats or mice have elevated rates of gluconeogenesis as do streptozotocin-induced diabetic rats or mice fed ad libitum. Rats are made diabetic with a single intravenous injection of 40 mg/kg of streptozotocin while C57BL/KsJ mice are given 40-60 mg/kg i.p. for 5 consecutive days. Blood glucose is measured from tail-tip blood and then compounds are administered via different routes (p.o., i.p., i.v., s.c.). Blood is collected at various times thereafter and glucose measured. Alternatively, compounds are administered for several days, then the animals are fasted overnight, blood is collected and plasma glucose measured. Compounds that inhibit glucose production will decrease plasma glucose levels compared to the vehicle-treated control group.

Insulin Sensitivity

Both ob/ob and db/db mice as well as diabetic Zucker rats are hyperglycemic, hyperinsulinemic and insulin resistant. The animals are pre-bled, their glucose levels measured, and then they are grouped so that the mean glucose level is the same for each group. Compounds are administered daily either q.d. or b.i.d. by different

routes (p.o., i.p., s.c.) for 7-28 days. Blood is collected at various times and plasma glucose and insulin levels determined. Compounds that improve insulin sensitivity in these models will decrease both plasma glucose and insulin levels when compared to the vehicle-treated control group.

Insulin Secretion

Compounds that enhance insulin secretion from the pancreas will increase plasma insulin levels and improve the disappearance of plasma glucose following the administration of a glucose load. When measuring insulin levels, compounds are administered by different routes (p.o., i.p., s.c. or i.v.) to overnight fasted normal rats or mice. At the appropriate time an intravenous glucose load (0.4g/kg) is given, blood is collected one minute later. Plasma insulin levels are determined. Compounds that enhance insulin secretion will increase plasma insulin levels compared to animals given only glucose. When measuring glucose disappearance, animals are bled at the appropriate time after compound administration, then given either an oral or intraperitoneal glucose load (1g/kg), bled again after 15, 30, 60 and 90 minutes and plasma glucose levels determined. Compounds that increase insulin levels will decrease glucose levels and the area-under-the glucose curve when compared to the vehicle-treated group given only glucose.

Compounds that enhance insulin secretion from the pancreas will increase plasma insulin levels and improve the disappearance of plasma glucose following the administration of a glucose load. When measuring insulin levels, test compounds which regulate receptor tyrosine phosphatase are administered by different routes (p.o., i.p., s.c., or i.v.) to overnight fasted normal rats or mice. At the appropriate time an intravenous glucose load (0.4g/kg) is given, blood is collected one minute later. Plasma insulin levels are determined. Test compounds that enhance insulin secretion will increase plasma insulin levels compared to animals given only glucose. When measuring glucose disappearance, animals are bled at the appropriate time after compound administration, then given either an oral or intraperitoneal glucose

load (1g/kg), bled again after 15, 30, 60, and 90 minutes and plasma glucose levels determined. Test compounds that increase insulin levels will decrease glucose levels and the area-under-the glucose curve when compared to the vehicle-treated group given only glucose.

Glucose Production

Over-production of glucose by the liver, due to an enhanced rate of gluconeogenesis, is the major cause of fasting hyperglycemia in diabetes. Overnight fasted normal rats or mice have elevated rates of gluconeogenesis as do streptozotocin-induced diabetic rats or mice fed ad libitum. Rats are made diabetic with a single intravenous injection of 40 mg/kg of streptozotocin while C57BL/KsJ mice are given 40-60 mg/kg i.p. for 5 consecutive days. Blood glucose is measured from tail-tip blood and then compounds are administered via different routes (p.o., i.p., i.v., s.c.). Blood is collected at various times thereafter and glucose measured. Alternatively, compounds are administered for several days, then the animals are fasted overnight, blood is collected and plasma glucose measured. Compounds that inhibit glucose production will decrease plasma glucose levels compared to the vehicle-treated control group.

Insulin Sensitivity

Both ob/ob and db/db mice as well as diabetic Zucker rats are hyperglycemic, hyperinsulinemic and insulin resistant. The animals are pre-bled, their glucose levels measured, and then they are grouped so that the mean glucose level is the same for each group. Compounds are administered daily either q.d. or b.i.d. by different routes (p.o., i.p., s.c.) for 7-28 days. Blood is collected at various times and plasma glucose and insulin levels determined. Compounds that improve insulin sensitivity in these models will decrease both plasma glucose and insulin levels when compared to the vehicle-treated control group.

Insulin Secretion

Compounds that enhance insulin secretion from the pancreas will increase plasma insulin levels and improve the disappearance of plasma glucose following the administration of a glucose load. When measuring insulin levels, compounds are administered by different routes (p.o., i.p., s.c. or i.v.) to overnight fasted normal rats or mice. At the appropriate time an intravenous glucose load (0.4g/kg) is given, blood is collected one minute later. Plasma insulin levels are determined. Compounds that enhance insulin secretion will increase plasma insulin levels compared to animals given only glucose. When measuring glucose disappearance, animals are bled at the appropriate time after compound administration, then given either an oral or intraperitoneal glucose load (1g/kg), bled again after 15, 30, 60 and 90 minutes and plasma glucose levels determined. Compounds that increase insulin levels will decrease glucose levels and the area-under-the glucose curve when compared to the vehicle-treated group given only glucose.

EXAMPLE 8

In vivo testing of compounds/target validation

Pain

Acute pain. Acute pain is measured on a hot plate mainly in rats. Two variants of hot plate testing are used: In the classical variant animals are put on a hot surface (52 to 56 °C) and the latency time is measured until the animals show nocifensive behavior, such as stepping or foot licking. The other variant is an increasing temperature hot plate where the experimental animals are put on a surface of neutral temperature. Subsequently this surface is slowly but constantly heated until the animals begin to lick a hind paw. The temperature which is reached when hind paw licking begins is a measure for pain threshold.

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v:, i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Persistent pain. Persistent pain is measured with the formalin or capsaicin test, mainly in rats. A solution of 1 to 5% formalin or 10 to 100 μg capsaicin is injected into one hind paw of the experimental animal. After formalin or capsaicin application the animals show nocifensive reactions like flinching, licking and biting of the affected paw. The number of nocifensive reactions within a time frame of up to 90 minutes is a measure for intensity of pain.

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to formalin or capsaicin administration.

Neuropathic pain. Neuropathic pain is induced by different variants of unilateral sciatic nerve injury mainly in rats. The operation is performed under anesthesia. The first variant of sciatic nerve injury is produced by placing loosely constrictive ligatures around the common sciatic nerve. The second variant is the tight ligation of about the half of the diameter of the common sciatic nerve. In the next variant, a group of models is used in which tight ligations or transections are made of either the L5 and L6 spinal nerves, or the L% spinal nerve only. The fourth variant involves an axotomy of two of the three terminal branches of the sciatic nerve (tibial and common peroneal nerves) leaving the remaining sural nerve intact whereas the last variant comprises the axotomy of only the tibial branch leaving the sural and common nerves uninjured. Control animals are treated with a sham operation.

Postoperatively, the nerve injured animals develop a chronic mechanical allodynia, cold allodynioa, as well as a thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA; Electronic von Frey

System, Somedic Sales AB, Hörby, Sweden). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy), or by means of a cold plate of 5 to 10 °C where the nocifensive reactions of the affected hind paw are counted as a measure of pain intensity. A further test for cold induced pain is the counting of nocifensive reactions, or duration of nocifensive responses after plantar administration of acetone to the affected hind limb. Chronic pain in general is assessed by registering the circadanian rhythms in activity (Surjo and Arndt, Universität zu Köln, Cologne, Germany), and by scoring differences in gait (foot print patterns; FOOTPRINTS program, Klapdor et al., 1997. A low cost method to analyze footprint patterns. J. Neurosci. Methods 75, 49-54).

Compounds are tested against sham operated and vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Inflammatory Pain. Inflammatory pain is induced mainly in rats by injection of 0.75 mg carrageenan or complete Freund's adjuvant into one hind paw. The animals develop an edema with mechanical allodynia as well as thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy, Paw thermal stimulator, G. Ozaki, University of California, USA). For edema measurement two methods are being used. In the first method, the animals are sacrificed and the affected hindpaws sectioned and weighed. The second method comprises differences in paw volume by measuring water displacement in a plethysmometer (Ugo Basile, Comerio, Italy).

Compounds are tested against uninflamed as well as vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Diabetic neuropathic pain. Rats treated with a single intraperitoneal injection of 50 to 80 mg/kg streptozotocin develop a profound hyperglycemia and mechanical allodynia within 1 to 3 weeks. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA).

Compounds are tested against diabetic and non-diabetic vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Parkinson's disease

6-Hydroxydopamine (6-OH-DA) Lesion.

Degeneration of the dopaminergic nigrostriatal and striatopallidal pathways is the central pathological event in Parkinson's disease. This disorder has been mimicked experimentally in rats using single/sequential unilateral stereotaxic injections of 6-OH-DA into the medium forebrain bundle (MFB).

Male Wistar rats (Harlan Winkelmann, Germany), weighing 200±250 g at the beginning of the experiment, are used. The rats are maintained in a temperature- and humidity-controlled environment under a 12 h light/dark cycle with free access to food and water when not in experimental sessions. The following in vivo protocols are approved by the governmental authorities. All efforts are made to minimize animal suffering, to reduce the number of animals used, and to utilize alternatives to in vivo techniques.

Animals are administered pargyline on the day of surgery (Sigma, St. Louis, MO, USA; 50 mg/kg i.p.) in order to inhibit metabolism of 6-OHDA by monoamine oxidase and desmethylimipramine HCl (Sigma; 25 mg/kg i.p.) in order to prevent

uptake of 6-OHDA by noradrenergic terminals. Thirty minutes later the rats are anesthetized with sodium pentobarbital (50 mg/kg) and placed in a stereotaxic frame. In order to lesion the DA nigrostriatal pathway 4 µl of 0.01% ascorbic acid-saline containing 8 µg of 6-OHDA HBr (Sigma) are injected into the left medial fore-brain bundle at a rate of 1 µl/min (2.4 mm anterior, 1.49 mm lateral, -2.7 mm ventral to Bregma and the skull surface). The needle is left in place an additional 5 min to allow diffusion to occur.

Stepping Test. Forelimb akinesia is assessed three weeks following lesion placement using a modified stepping test protocol. In brief, the animals are held by the experimenter with one hand fixing the hindlimbs and slightly raising the hind part above the surface. One paw is touching the table, and is then moved slowly sideways (5 s for 1 m), first in the forehand and then in the backhand direction. The number of adjusting steps is counted for both paws in the backhand and forehand direction of movement. The sequence of testing is right paw forehand and backhand adjusting stepping, followed by left paw forehand and backhand directions. The test is repeated three times on three consecutive days, after an initial training period of three days prior to the first testing. Forehand adjusted stepping reveals no consistent differences between lesioned and healthy control animals. Analysis is therefore restricted to backhand adjusted stepping.

Balance Test. Balance adjustments following postural challenge are also measured during the stepping test sessions. The rats are held in the same position as described in the stepping test and, instead of being moved sideways, tilted by the experimenter towards the side of the paw touching the table. This maneuver results in loss of balance and the ability of the rats to regain balance by forelimb movements is scored on a scale ranging from 0 to 3. Score 0 is given for a normal forelimb placement. When the forelimb movement is delayed but recovery of postural balance detected, score 1 is given. Score 2 represents a clear, yet insufficient, forelimb reaction, as evidenced by muscle contraction, but lack of success in recovering balance, and score 3 is given for no reaction of movement. The test is repeated three times a day on each

side for three consecutive days after an initial training period of three days prior to the first testing.

Staircase Test (Paw Reaching). A modified version of the staircase test is used for evaluation of paw reaching behavior three weeks following primary and secondary lesion placement. Plexiglass test boxes with a central platform and a removable staircase on each side are used. The apparatus is designed such that only the paw on the same side at each staircase can be used, thus providing a measure of independent forelimb use. For each test the animals are left in the test boxes for 15 min. The double staircase is filled with 7 x 3 chow pellets (Precision food pellets, formula: P, purified rodent diet, size 45 mg; Sandown Scientific) on each side. After each test the number of pellets eaten (successfully retrieved pellets) and the number of pellets taken (touched but dropped) for each paw and the success rate (pellets eaten/pellets taken) are counted separately. After three days of food deprivation (12 g per animal per day) the animals are tested for 11 days. Full analysis is conducted only for the last five days.

MPTP treatment. The neurotoxin 1-methyl-4-phenyl-1,2,3,6- tetrahydro-pyridine (MPTP) causes degeneration of mesencephalic dopaminergic (DAergic) neurons in rodents, non-human primates, and humans and, in so doing, reproduces many of the symptoms of Parkinson's disease. MPTP leads to a marked decrease in the levels of dopamine and its metabolites, and in the number of dopaminergic terminals in the striatum as well as severe loss of the tyrosine hydroxylase (TH)-immunoreactive cell bodies in the substantia nigra, pars compacta.

In order to obtain severe and long-lasting lesions, and to reduce mortality, animals receive single injections of MPTP, and are then tested for severity of lesion 7–10 days later. Successive MPTP injections are administered on days 1, 2 and 3. Animals receive application of 4 mg/kg MPTP hydrochloride (Sigma) in saline once daily. All injections are intraperitoneal (i.p.) and the MPTP stock solution is frozen between injections. Animals are decapitated on day 11.

Immunohistology. At the completion of behavioral experiments, all animals are anaesthetized with 3 ml thiopental (1 g/40 ml i.p., Tyrol Pharma). The mice are perfused transcardially with 0.01 M PBS (pH 7.4) for 2 min, followed by 4% paraformaldehyde (Merck) in PBS for 15 min. The brains are removed and placed in 4% paraformaldehyde for 24 h at 4 °C. For dehydration they are then transferred to a 20% sucrose (Merck) solution in 0.1 M PBS at 4 °C until they sink. The brains are frozen in methylbutan at -20 °C for 2 min and stored at -70 °C. Using a sledge microtome (mod. 3800-Frigocut, Leica), 25 μm sections are taken from the genu of the corpus callosum (AP 1.7 mm) to the hippocampus (AP 21.8 mm) and from AP 24.16 to AP 26.72. Forty-six sections are cut and stored in assorters in 0.25 M Tris buffer (pH 7.4) for immunohistochemistry.

A series of sections is processed for free-floating tyrosine hydroxylase (TH) immunohistochemistry. Following three rinses in 0.1 M PBS, endogenous peroxidase activity is quenched for 10 min in 0.3% H_2O_2 ±PBS. After rinsing in PBS, sections are preincubated in 10% normal bovine serum (Sigma) for 5 min as blocking agent and transferred to either primary anti-rat TH rabbit antiserum (dilution 1:2000).

Following overnight incubation at room temperature, sections for TH immunoreactivity are rinsed in PBS (2 x10 min) and incubated in biotinylated anti-rabbit immunoglobulin G raised in goat (dilution 1:200) (Vector) for 90 min, rinsed repeatedly and transferred to Vectastain ABC (Vector) solution for 1 h. 3,.3' -Diaminobenzidine tetrahydrochloride (DAB; Sigma) in 0.1 M PBS, supplemented with 0.005% H_2O_2 , serves as chromogen in the subsequent visualization reaction. Sections are mounted on to gelatin-coated slides, left to dry overnight, counter-stained with hematoxylin dehydrated in ascending alcohol concentrations and cleared in butylacetate. Coverslips are mounted on entellan.

Rotarod Test. We use a modification of the procedure described by Rozas and Labandeira-Garcia (1997), with a CR-1 Rotamex system (Columbus Instruments,

Columbus, OH) comprising an IBM-compatible personal computer, a CIO-24 data acquisition card, a control unit, and a four-lane rotarod unit. The rotarod unit consists of a rotating spindle (diameter 7.3 cm) and individual compartments for each mouse. The system software allows preprogramming of session protocols with varying rotational speeds (0–80 rpm). Infrared beams are used to detect when a mouse has fallen onto the base grid beneath the rotarod. The system logs the fall as the end of the experiment for that mouse, and the total time on the rotarod, as well as the time of the fall and all the set-up parameters, are recorded. The system also allows a weak current to be passed through the base grid, to aid training.

Dementia

The object recognition task. The object recognition task has been designed to assess the effects of experimental manipulations on the cognitive performance of rodents. A rat is placed in an open field, in which two identical objects are present. The rats inspects both objects during the first trial of the object recognition task. In a second trial, after a retention interval of for example 24 hours, one of the two objects used in the first trial, the 'familiar' object, and a novel object are placed in the open field. The inspection time at each of the objects is registered. The basic measures in the OR task is the time spent by a rat exploring the two object the second trial. Good retention is reflected by higher exploration times towards the novel than the 'familiar' object.

Administration of the putative cognition enhancer prior to the first trial predominantly allows assessment of the effects on acquisition, and eventually on consolidation processes. Administration of the testing compound after the first trial allows to assess the effects on consolidation processes, whereas administration before the second trial allows to measure effects on retrieval processes.

The passive avoidance task. The passive avoidance task assesses memory performance in rats and mice. The inhibitory avoidance apparatus consists of a

two-compartment box with a light compartment and a dark compartment. The two compartments are separated by a guillotine door that can be operated by the experimenter. A threshold of 2 cm separates the two compartments when the guillotine door is raised. When the door is open, the illumination in the dark compartment is about 2 lux. The light intensity is about 500 lux at the center of the floor of the light compartment.

Two habituation sessions, one shock session, and a retention session are given, separated by inter-session intervals of 24 hours. In the habituation sessions and the retention session the rat is allowed to explore the apparatus for 300 sec. The rat is placed in the light compartment, facing the wall opposite to the guillotine door. After an accommodation period of 15 sec. the guillotine door is opened so that all parts of the apparatus can be visited freely. Rats normally avoid brightly lit areas and will enter the dark compartment within a few seconds.

In the shock session the guillotine door between the compartments is lowered as soon as the rat has entered the dark compartment with its four paws, and a scrambled 1 mA footshock is administered for 2 sec. The rat is removed from the apparatus and put back into its home cage. The procedure during the retention session is identical to that of the habituation sessions.

The step-through latency, that is the first latency of entering the dark compartment (in sec.) during the retention session is an index of the memory performance of the animal; the longer the latency to enter the dark compartment, the better the retention is. A testing compound in given half an hour before the shock session, together with 1 $mg*kg^{-1}$ scopolamine. Scopolamine impairs the memory performance during the retention session 24 hours later. If the test compound increases the enter latency compared with the scopolamine-treated controls, is likely to possess cognition enhancing potential.

The Morris water escape task. The Morris water escape task measures spatial orientation learning in rodents. It is a test system that has extensively been used to investigate the effects of putative therapeutic on the cognitive functions of rats and mice. The performance of an animal is assessed in a circular water tank with an escape platform that is submerged about 1 cm below the surface of the water. The escape platform is not visible for an animal swimming in the water tank. Abundant extra-maze cues are provided by the furniture in the room, including desks, computer equipment, a second water tank, the presence of the experimenter, and by a radio on a shelf that is playing softly.

The animals receive four trials during five daily acquisition sessions. A trial is started by placing an animal into the pool, facing the wall of the tank. Each of four starting positions in the quadrants north, east, south, and west is used once in a series of four trials; their order is randomized. The escape platform is always in the same position. A trial is terminated as soon as the animal had climbs onto the escape platform or when 90 seconds have elapsed, whichever event occurs first. The animal is allowed to stay on the platform for 30 seconds. Then it is taken from the platform and the next trial is started. If an animal did not find the platform within 90 seconds it is put on the platform by the experimenter and is allowed to stay there for 30 seconds. After the fourth trial of the fifth daily session, an additional trial is given as a probe trial: the platform is removed, and the time the animal spends in the four quadrants is measured for 30 or 60 seconds. In the probe trial, all animals start from the same start position, opposite to the quadrant where the escape platform had been positioned during acquisition.

Four different measures are taken to evaluate the performance of an animal during acquisition training: escape latency, traveled distance, distance to platform, and swimming speed. The following measures are evaluated for the probe trial: time (s) in quadrants and traveled distance (cm) in the four quadrants. The probe trial provides additional information about how well an animal learned the position of the escape platform. If an animal spends more time and swims a longer distance in the

quadrant where the platform had been positioned during the acquisition sessions than in any other quadrant, one concludes that the platform position has been learned well.

In order to assess the effects of putative cognition enhancing compounds, rats or mice with specific brain lesions which impair cognitive functions, or animals treated with compounds such as scopolamine or MK-801, which interfere with normal learning, or aged animals which suffer from cognitive deficits, are used.

The T-maze spontaneous alternation task. The T-maze spontaneous alternation task (TeMCAT) assesses the spatial memory performance in mice. The start arm and the two goal arms of the T-maze are provided with guillotine doors which can be operated manually by the experimenter. A mouse is put into the start arm at the beginning of training. The guillotine door is closed. In the first trial, the 'forced trial', either the left or right goal arm is blocked by lowering the guillotine door. After the mouse has been released from the start arm, it will negotiate the maze, eventually enter the open goal arm, and return to the start position, where it will be confined for 5 seconds, by lowering the guillotine door. Then, the animal can choose freely between the left and right goal arm (all guillotine-doors opened) during 14 'free choice' trials. As soon a the mouse has entered one goal arm, the other one is closed. The mouse eventually returns to the start arm and is free to visit whichever go alarm it wants after having been confined to the start arm for 5 seconds. After completion of 14 free choice trials in one session, the animal is removed from the maze. During training, the animal is never handled.

The percent alternations out of 14 trials is calculated. This percentage and the total time needed to complete the first forced trial and the subsequent 14 free choice trials (in s) is analyzed. Cognitive deficits are usually induced by an injection of scopolamine, 30 min before the start of the training session. Scopolamine reduced the per-cent alternations to chance level, or below. A cognition enhancer, which is always administered before the training session, will at least partially, antagonize the scopolamine-induced reduction in the spontaneous alternation rate.

EXAMPLE 9

Identification of test compound efficacy in a COPD animal model

Guinea pigs are exposed on a single occasion to tobacco smoke for 50 minutes. Animals are sacrificed between 10 minutes and 24 hour following the end of the exposure and their lungs placed in RNAlater™. The lung tissue is homogenised ,and total RNA was extracted using a Qiagens RNeasy™ Maxi kit. Molecular Probes RiboGreen™ RNA quantitation method is used to quantify the amount of RNA in each sample.

Total RNA is reverse transcribed, and the resultant cDNA is used in a real-time polymerase chain reaction (PCR). The cDNA is added to a solution containing the sense and anti-sense primers and the 6-carboxy-tetramethyl-rhodamine labeled probe of the receptor tyrosine phosphatase gene. Cyclophilin is used as the housekeeping gene. The expression of the receptor tyrosine phosphatase gene is measured using the TaqMan real-time PCR system that generates an amplification curve for each sample. From this curve a threshold cycle value is calculated: the fractional cycle number at which the amount of amplified target reaches a fixed threshold. A sample containing many copies of the receptor tyrosine phosphatase gene will reach this threshold earlier than a sample containing fewer copies. The threshold is set at 0.2, and the threshold cycle C_T is calculated from the amplification curve. The C_T value for the receptor tyrosine phosphatase gene is normalized using the C_T value for the housekeeping gene.

Expression of the receptor tyrosine phosphatase gene is increased by at least 3-fold between 10 minutes and 3 hours post tobacco smoke exposure compared to air exposed control animals.

Test compounds are evaluated as follows. Animals are pre-treated with a test compound between 5 minutes and 1 hour prior to the tobacco smoke exposure and they are then sacrificed up to 3 hours after the tobacco smoke exposure has been completed. Control animals are pre-treated with the vehicle of the test compound via the route of administration chosen for the test compound. A test compound that reduces the tobacco smoke induced upregulation of receptor tyrosine phosphatase gene relative to the expression seen in vehicle treated tobacco smoke exposed animals is identified as an inhibitor of receptor tyrosine phosphatase gene expression.

REFERENCES

1. Wright MB, Hugo C, Seifert R, Disteche CM, Bowen-Pope DF. Proliferating and migrating mesangial cells responding to injury express a novel receptor protein-tyrosine phosphatase in experimental mesangial proliferative glomerulonephritis. J Biol Chem 1998 Sep 11;273(37):23929-37.

2. Zelivianski S, Dean J, Madhavan D, Lin FF, Lin MF. Expression of receptor protein tyrosine phosphatase alpha mRNA in human prostate cancer cell lines. Mol Cell Biochem 2000 May;208(1-2):11-8.

3. Fauman EB, Saper MA. Structure and function of the protein tyrosine phosphatases. Trends Biochem Sci 1996 Nov;21(11):413-7.

4. Neel BG, Tonks NK. Protein tyrosine phosphatases in signal transduction. Curr Opin Cell Biol 1997 Apr;9(2):193-204.

5. Calvert-Evers JL, Hammond KD. Temporal variations in protein tyrosine phosphatase activity during cell proliferation and differentiation. Cell Biol Int 2000;24(8):559-68.

CLAIMS

1. An isolated polynucleotide being selected from the group consisting of:
   a. a polynucleotide encoding a receptor tyrosine phosphatase polypeptide comprising an amino acid sequence selected form the group consisting of:
      i. amino acid sequences which are at least about 87% identical to the amino acid sequence shown in SEQ ID NO: 2; and
      ii. the amino acid sequence shown in SEQ ID NO: 2.
   b. a polynucleotide comprising the sequence of SEQ ID NO: 1;
   c. a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a receptor tyrosine phosphatase polypeptide;
   d. a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a receptor tyrosine phosphatase polypeptide; and
   e. a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a receptor tyrosine phosphatase polypeptide.

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified receptor tyrosine phosphatase polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a receptor tyrosine phosphatase polypeptide, wherein the method comprises the following steps:

a. culturing the host cell of claim 3 under conditions suitable for the expression of the receptor tyrosine phosphatase polypeptide; and

b. recovering the receptor tyrosine phosphatase polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a receptor tyrosine phosphatase polypeptide in a biological sample comprising the following steps:

a. hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b. detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a receptor tyrosine phosphatase polypeptide of claim 4 comprising the steps of:

a. contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the receptor tyrosine phosphatase polypeptide and

b. detecting the interaction

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a receptor tyrosine phosphatase, comprising the steps of:

a. contacting a test compound with any receptor tyrosine phosphatase polypeptide encoded by any polynucleotide of claim1;

b. detecting binding of the test compound to the receptor tyrosine phosphatase polypeptide, wherein a test compound which binds to the

polypeptide is identified as a potential therapeutic agent for decreasing the activity of a receptor tyrosine phosphatase.

11. A method of screening for agents which regulate the activity of a receptor tyrosine phosphatase, comprising the steps of:
    a. contacting a test compound with a receptor tyrosine phosphatase polypeptide encoded by any polynucleotide of claim 1; and
    b. detecting a receptor tyrosine phosphatase activity of the polypeptide, wherein a test compound which increases the receptor tyrosine phosphatase activity is identified as a potential therapeutic agent for increasing the activity of the receptor tyrosine phosphatase, and wherein a test compound which decreases the receptor tyrosine phosphatase activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the receptor tyrosine phosphatase.

12. A method of screening for agents which decrease the activity of a receptor tyrosine phosphatase, comprising the step of:

    contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of receptor tyrosine phosphatase.

13. A method of reducing the activity of receptor tyrosine phosphatase, comprising the step of:

    contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any receptor tyrosine phosphatase polypeptide of claim 4, whereby the activity of receptor tyrosine phosphatase is reduced.

14. A reagent that modulates the activity of a receptor tyrosine phosphatase polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15. A pharmaceutical composition, comprising:

the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a receptor tyrosine phosphatase in a disease.

17. Use of claim 16 wherein the disease is diabetes, a CNS disorder, obesity, COPD or a cardiovascular disorder.

Fig. 1

```
atggattttc ttatcatttt tcttttactt tttattggga cttcagagac acaggttgat
gtttccaatg tcgttcctgg tactaggtac gatataacca tctcttcaat ttctacaaca
tacacctcac ctgttactag aatagtgaca acaaatgtaa caaaaccagg gcctccagtc
ttcctagccg gggaaagagt cggatctgct gggattcttc tgtcttggaa tacaccacct
aatccaaatg gaaggattat atcttacatt gtcaaatata aggaagtttg tccgtggatg
caaacagtat atacacaagt cagatcaaag ccagacagtc tggaagttct tcttactaat
cttaatcctg gaacaacata tgaaattaag gttgctgctg aaaacagtgc tggcattgga
gtgtttagtg atccatttct cttccaaact gcagaaagtg ctccaggaaa agtggtgaat
ctcacagttg aggcctacaa cgcttcagca gttaagctga tttggtattt acctcggcaa
ccaaatggca aaattaccag cttcaagatt agtgtcaagc atgccagaag tgggatagta
gtgaaagatg tctcaatcag agtagaggac attttgactg ggaaattgcc agaaagcaat
gagaatagtg aatctttttt atggagtaca gccagcccct ctccaaccct tggtagagtt
acacctccat cgcgtaccac acattcatca agcacgttga cacagaatga gatcagctct
gtgtggaaag agcctatcag ttttgtagtg acacacttga gaccttatac aacatatctt
tttgaagttt cagctgctac aactgaagca ggttatattg atagtacgat tgtcagaaca
ccagaatcag tgcctgaagg accaccacaa aactgcgtaa caggcaacat cacaggaaag
tccttttcaa ttttatggga cccaccaact atagtaacag ggaaatttag ttatagagtt
gaattatatg gaccatcagg tcgcattttg gataacagca caaaagacct caagtttgca
ttcactaacc taacaccatt tacaatgtat gatgtctata ttgcggctga aaccagtgca
gggactgggc ccaagtcaaa tatttcagta ttcactccac cagatgttcc aggggcagtg
tttgatttac aacttgcaga ggtagaatcc acgcaagtaa gaattacttg gaagaaacca
cgacaaccaa atggaattat taaccaatac cgagtgaaag tgctagttcc agagacagga
ataattttgg aaaatacttt gctcactgga aataatgagt atataaatga ccccatggct
ccagaaattg tgaacatagt agagccaatg gtaggattat atgagggttc agcagagatg
tcgtctgacc ttcactcact tgctacattt atatataaca gccatccaga taaaaacttt
cctgcaagga atagagctga agaccagact tcaccagttg taactacaag gaatcagtat
attactgaca ttgcagctga acagctgtct tatgttatca ggagacttgt acctttcact
gagcacatga ttagtgtatc tgctttcacc atcatgggag aaggaccacc aacagttctc
agtgttagga cacgtcagca agtgccaagc tccattaaaa ttataaacta taaaaatatt
agttcttcat ctattttgtt atattgggat cctccagaat atcccaatgg aaaaataact
cactatacga tttatgcaat ggaattggat acaaacagag cattccagat aactaccata
gataacagct ttctcataac agggttaaag aaatacacaa aatacaaaat gagagtggca
gcctcaaccc acgttggaga aagttctttg tctgaagaaa atgacatctt tgtgagaact
tcagaagatg aaccggaatc atcacctcaa gatgtcgaag taattgatgt taccgcagat
gaaataaggt tgaagtggtc accacccgaa aagcccaatg ggatcattat tgcttatgaa
gtgctatata aaaatataga tactttatat atgaagaaca catcaacaac agacataata
ttaaggaact taagacctca caccctctat aacatttctg taaggtctta caccagattt
ggtcatggca atcaggtatc ttctttactc tctgtaagga cttcggagac tgtgcctgat
agtgcaccag aaaatatcac ttacaaaaat atttcttctg gagagattga gctatcattc
cttcccccaa gtagtcccaa tggaatcata caaaaatata caatttatct caagagaagt
aatggaaatg aggaaagaac tataaataca acctctttaa cccaaaacat taaagtactg
aagaaatata cccaatatat cattgaggtg tctgctagta cactcaaagg tgaaggagtt
cggagtgctc ccataagtat actgacggag gaagatgctc ctgattctcc ccctcaagac
ttctctgtaa aacagttgtc tggtgtcacg gtgaagttgt catggcaacc acccctggag
ccaaatggaa ttatccttta ttacacagtt tatgtctgga atagatcatc attaaaaact
attaatgtca ctgaaacatc attggagtta tcagatttgg attataatgt tgaatacagt
gcttatgtaa cagctagcac cagatttggt gatgggaaaa caagaagcaa tatcattagc
tttcaaacac cagagggagc accaagcgat cctcccaaag atgtttatta tgcaaacctc
agttcttcat caataattct tttctggaca cctccttcaa aacctaatgg gattatacaa
tattactctg tttattacag aaatacttca ggtactttta tgcagaattt tacactccat
gaagtaacca atgactttga caatatgact gtatccacaa ttatagataa actgacaata
ttcagctact atacattttg gttaacagca agtacttcag ttggaaatgg gaataaaagc
agtgacatca ttgaagtata cacagatcaa gacatacctg aagggtttgt tggaaacctg
acttacgaat ccatttcgtc aactgcaata aatgtaagct gggtcccacc ggctcaacca
aacggtctag tcttctacta tgtttcactg atcttacagc agactcctcg ccatgtgaga
ccacctcttg ttacatatga gagaagcata tattttgata atctggaaaa atacactgat
tatatattaa aaattactcc atcaacagaa aagggattct ctgataccta tactgcccag
ctatacatca agactgaaga agatgtccca gaaacttcac caataatcaa cacttttaaa
```

Fig. 1 (continued)

aacctttcct ctacctcagt tctcttatca tgggatcccc cagtaaagcc aaatggtgca
ataataagtt atgatttaac tttacaagga ccaaatgaaa attattcttt cattacttct
gataattaca taatattgga agagctttca ccatttacat tatatagctt ttttgctgcc
gcaagaacta gaaaaggact tggtccttcc agtattcttt tcttttacac agatgagtca
gtgccgttag cacctccaca aaatttgact ttaatcaact gtacttcaga ctttgtatgg
ctgaaatgga gcccaagtcc tcttccaggt ggtattgtta aagtatatag ttttaaaatt
catgaacatg aaactgacac tatatattat aagaatatat caggatttaa aactgaagcc
aaacttgttg gactggaacc agtcagcacc tactctatcc gtgtatctgc gttcaccaaa
gttggaaatg gcaatcaatt tagtaatgta gtaaaattca caacccaaga atcagttcca
gatgtcgtgc agaatatgca gtgcatggca actagctggc agtcagtttt agtgaaatgg
gatccaccca aaaaggcaaa tggaataata acgcagtata tggtaacagt tgaaaggaat
tctacaaaag tttctcccca agatcacatg tacactttca taaagcttct tgccaatacc
tcatatgtct ttaaagtaag agcttcaacc tcagctggtg aaggtgatga aagcacatgc
catgtcagca cactacctga aacagttccc agtgttccca caaatattgc tttttctgat
gttcagtcaa ctagtgcaac attgacatgg ataagacctg acactatcct tggctacttt
caaaattaca aaattaccac tcaacttcgt gctcaaaaat gcaaagaatg ggaatccgaa
gaatgtgttg aatatcaaaa aattcaatac ctctatgaag ctcacttaac tgaagagaca
gtatatggat taaagaaatt tagatggtat agattccaag tggctgccag caccaatgct
ggctatggca atgcttcaaa ctggatttct acaaaaactc tgcctggccc tccagatggt
cctcctgaaa atgttcatgt agtagcaaca tcacctttta gcatcagcat aagctggagt
gaacctgctg tcattactgg accaacatgt tatctgattg atgtcaaatc ggtagataat
gatgaattta atatatcctt catcaagtca aatgaagaaa ataaaaccat agaaattaaa
gatttagaaa tattcacaag gtattctgta gtgatcactg catttactgg gaacattagt
gctgcatatg tagaagggaa gtcaagtgct gaaatgattg ttactacttt agaatcagcc
ccaaaggacc cacctaacaa catgacattt cagaagatac cagatgaagt tacaaaattt
caattaacgt tccttcctcc ttctcaacct aatggaaata tccaagtata tcaagctctg
gtttaccgag aagatgatcc tactgctgtc cagattcaca acctcagtat tatacagaaa
accaacacat tcgtcattgc aatgctagaa ggactaaaag gtggacatac atacaatatc
agtgtttacg cagtcaatag tgctggtgca ggtccaaagg ttccgatgag aataaccatg
gatatcaaag ctccagcacg accaaaaacc aaaccaaccc ctatttatga tgccacagga
aaactgcttg tgacttcaac aacaattaca atcagaatgc caatatgtta ctacagtgat
gatcatggac caataaaaaa tgtacaagtg cttgtgacag aaacaggagc tcagcatgat
ggaaatgtaa caaagtggta tgatgcatat tttaataaag caaggccata ttttacaaat
gaaggctttc ctaaccctcc atgtacagaa ggaaagacaa agtttagtgg caatgaagaa
atctacatca taggtgctga taatgcatgc atgattcctg gcaatgaaga caaaatttgc
aatggaccac tgaaaccaaa aaagcaatac ttatttaaat ttagagctac aaatattatg
ggacaattta ctgactctga ttattctgac cctgttaaga ctttagggga aggactttca
gaaagaaccg tagagatcat tctttccgtc actttgtgta tcctttcaat aattctcctt
ggaacagcta tttttgcatt tgcaagaatt cgacagaagc agaaagaagg tggcacatac
tctcctcagg atgcagaaat tattgacact aaattgaagc tggatcagct catcacagtg
gcagacctgg aactgaagga cgagagatta acgcggttac ttagttatag aaaatccatc
aagccaataa gcaagaaatc cttcctgcaa catgttgaag agctttgcac aaacaacaac
ctaaagtttc aagaagaatt ttcggaatta ccaaaatttc ttcaggatct ttcttcaact
gatgctgatc tgccttggaa tagagcaaaa aaccgcttcc caaacataaa accatataat
aataacagag taaagctgat agctgacgct agtgttccag gttcggatta tattaatgcc
agctatattt ctggttattt atgtccaaat gaatttattg ctactcaagg tccactacca
ggaacagttg gagatttttg gaaaatggtg tgggaaacca gagcaaaaac attagtaatg
ctaacacagt gttttgaaaa aggacggatc agatgccatc agtattggcc agaggacaac
aagccagtta ctgtctttgg agatatagtg attacaaagc taatggagga tgttcaaata
gattggacta tcagggatct gaaaattgaa aggcatgggg attgcatgac tgttcgacag
tgtaacttta ctgcctggcc agagcatggg gttcctgaga acagcgcccc tctaattcac
tttgtgaagt tggttcgagc aagcagggca catgacacca cacctatgat tgttcactgc
agtgctggag ttggaagaac tggagttttt attgctctgg accatttaac acaacatata
aatgaccatg attttgtgga tatatatgga ctagtagctg aactgagaag tgaaagaatg
tgcatggtgc agaatctggc acagtatatc tttttacacc agtgcattct ggatctctta
tcaaataagg gaagtaatca gcccatctgt tttgttaact attcagcact tcagaagatg
gactctttgg acgccatgga aggtgatgtt gagcttgaat gggaagaaac cactatg

Fig. 2

MDFLIIFLLL FIGTSETQVD VSNVVPGTRY DITISSISTT YTSPVTRIVT TNVTKPGPPV
FLAGERVGSA GILLSWNTPP NPNGRIISYI VKYKEVCPWM QTVYTQVRSK PDSLEVLLTN
LNPGTTYEIK VAAENSAGIG VFSDPFLFQT AESAPGKVVN LTVEAYNASA VKLIWYLPRQ
PNGKITSFKI SVKHARSGIV VKDVSIRVED ILTGKLPESN ENSESFLWST ASPSPTLGRV
TPPSRTTHSS STLTQNEISS VWKEPISFVV THLRPYTTYL FEVSAATTEA GYIDSTIVRT
PESVPEGPPQ NCVTGNITGK SFSILWDPPT IVTGKFSYRV ELYGPSGRIL DNSTKDLKFA
FTNLTPFTMY DVYIAAETSA GTGPKSNISV FTPPDVPGAV FDLQLAEVES TQVRITWKKP
RQPNGIINQY RVKVLVPETG IILENTLLTG NNEYINDPMA PEIVNIVEPM VGLYEGSAEM
SSDLHSLATF IYNSHPDKNF PARNRAEDQT SPVVTTRNQY ITDIAAEQLS YVIRRLVPFT
EHMISVSAFT IMGEGPPTVL SVRTRQQVPS SIKIINYKNI SSSSILLYWD PPEYPNGKIT
HYTIYAMELD TNRAFQITTI DNSFLITGLK KYTKYKMRVA ASTHVGESSL SEENDIFVRT
SEDEPESSPQ DVEVIDVTAD EIRLKWSPPE KPNGIIIAYE VLYKNIDTLY MKNTSTTDII
LRNLRPHTLY NISVRSYTRF GHGNQVSSLL SVRTSETVPD SAPENITYKN ISSGEIELSF
LPPSSPNGII QKYTIYLKRS NGNEERTINT TSLTQNIKVL KKYTQYIIEV SASTLKGEGV
RSAPISILTE EDAPDSPPQD FSVKQLSGVT VKLSWQPPLE PNGIILYYTV YVWNRSSLKT
INVTETSLEL SDLDYNVEYS AYVTASTRFG DGKTRSNIIS FQTPEGAPSD PPKDVYYANL
SSSSIILFWT PPSKPNGIIQ YYSVYYRNTS GTFMQNFTLH EVTNDFDNMT VSTIIDKLTI
FSYYTFWLTA STSVGNGNKS SDIIEVYTDQ DIPEGFVGNL TYESISSTAI NVSWVPPAQP
NGLVFYYVSL ILQQTPRHVR PPLVTYERSI YFDNLEKYTD YILKITPSTE KGFSDTYTAQ
LYIKTEEDVP ETSPIINTFK NLSSTSVLLS WDPPVKPNGA IISYDLTLQG PNENYSFITS
DNYIILEELS PFTLYSFFAA ARTRKGLGPS SILFFYTDES VPLAPPQNLT LINCTSDFVW
LKWSPSPLPG GIVKVYSFKI HEHETDTIYY KNISGFKTEA KLVGLEPVST YSIRVSAFTK
VGNGNQFSNV VKFTTQESVP DVVQNMQCMA TSWQSVLVKW DPPKKANGII TQYMVTVERN
STKVSPQDHM YTFIKLLANT SYVFKVRAST SAGEGDESTC HVSTLPETVP SVPTNIAFSD
VQSTSATLTW IRPDTILGYF QNYKITTQLR AQKCKEWESE ECVEYQKIQY LYEAHLTEET
VYGLKKFRWY RFQVAASTNA GYGNASNWIS TKTLPGPPDG PPENVHVVAT SPFSISISWS
EPAVITGPTC YLIDVKSVDN DEFNISFIKS NEENKTIEIK DLEIFTRYSV VITAFTGNIS
AAYVEGKSSA EMIVTTLESA PKDPPNNMTF QKIPDEVTKF QLTFLPPSQP NGNIQVYQAL
VYREDDPTAV QIHNLSIIQK TNTFVIAMLE GLKGGHTYNI SVYAVNSAGA GPKVPMRITM
DIKAPARPKT KPTPIYDATG KLLVTSTTIT IRMPICYYSD DHGPIKNVQV LVTETGAQHD
GNVTKWYDAY FNKARPYFTN EGFPNPPCTE GKTKFSGNEE IYIIGADNAC MIPGNEDKIC
NGPLKPKKQY LFKFRATNIM GQFTDSDYSD PVKTLGEGLS ERTVEIILSV TLCILSIILL
GTAIFAFARI RQKQKEGGTY SPQDAEIIDT KLKLDQLITV ADLELKDERL TRLLSYRKSI
KPISKKSFLQ HVEELCTNNN LKFQEEFSEL PKFLQDLSST DADLPWNRAK NRFPNIKPYN
NNRVKLIADA SVPGSDYINA SYISGYLCPN EFIATQGPLP GTVGDFWKMV WETRAKTLVM
LTQCFEKGRI RCHQYWPEDN KPVTVFGDIV ITKLMEDVQI DWTIRDLKIE RHGDCMTVRQ
CNFTAWPEHG VPENSAPLIH FVKLVRASRA HDTTPMIVHC SAGVGRTGVF IALDHLTQHI
NDHDFVDIYG LVAELRSERM CMVQNLAQYI FLHQCILDLL SNKGSNQPIC FVNYSALQKM
DSLDAMEGDV ELEWEETTM

Fig. 3

MMDFHFSFLFLLIGTSESQVDVSSSFDGTGYDITLSSVSATTYSS
PVSRTLATNVTKPGPPVFLAGERVGSAGILLSWNTPPNPNGRIISYVVKYKEVCPWMQT
AYTRARAKPDSLEVLLTNLNPGTTYEIKVAAENNAGIGVFSDPFLFQTAESAPGKVVNL
TVEALNYSAVNLIWYLPRQPNGKITSFKISVKHARSGIVVKDVSLRVEDILSGKLPECN
ENSESFLWSTTSPSPTLGRVTPTVRTTQSSSTAARSKISSVWKEPISFVVTHLRPYTTY
LFEVSAVTTEAGYIDSTIVRTPESVPEGPPQNCIMGNVTGKAFSISWDPPTIVTGKFSY
RVELYGPSGRILDNSTKDLRFAFTHLTPFTMYDVYVAAETSAGVGPKSNLSVFTPPDVP
GAVFDLQIAEVEATEIRITWRKPRQPNGIISQYRVKVSVLETGVVLENTLLTGQDESIS
NPMSPEIMNLVDPMIGFYEGSGEMSSDLHSPASFIYNSHPHNDFPASTRAEEQSSPVVT
TRNQYMTDITAEQLSYVVRRLVPFTEHTISVSAFTIMGEGPPTVLTVRTREQVPSSIQI
INYKNISSSSILLYWDPPEYPNGKITHYTIYATELDTNRAFQMTTVDNSFLITGLKKYT
RYKMRVAASTHVGESSLSEENDIFVRTPEDEPESSPQDVQVTGVSPSELRLKWSPPEKP
NGIIIAYEVLYQNADTLFVKNTSTTDIIISDLKPYTLYNISIRSYTRLGHGNQSSSLLS
VRTSETVPDSAPENITYKNISSGEIEISFLPPRSPNGIIQKYTIYLKRSNSHEARTINT
TSLTQTIGGLKKYTHYVIEVSASTLKGEGIRSRPISILTEEDAPDSPPQNFSVKQLSGV
TVMLSWQPPLEPNGIILYYTVYVWDKSSLRAINATEASLVLSDLDYNVDYGACVTASTR
FGDGNARSSIINFRTPEGEPSDPPNDVHYVNLSSSSIILFWTPPVKPNGIIQYYSVYYQ
NTSGTFVQNFTLLQVTKESDNVTVSARIYRLAIFSYYTFWLTASTSVGNGNKSSDIIHV
YTDQDIPEGPVGNLTFESISSTAIHVSWEPPSQPNGLVFYYLSLNLQQSPPRHMIPPLV
TYENSIDFDDLEKYTDYIFKITPSTEKGFSETYTTQLHIKTEEDVPDTPPIINTFKNLS
STSILLSWDPPLKPNGAILGYHLTLQGPHANHTFVTSGNHIVLEELSPFTLYSFFAAAR
TMKGLGPSSILFFYTDESAPLAPPQNLTLINYTSDFVWLTWSPSPLPGGIVKVYSFKIH
EHETDTVFYKNISGLQTDAKLEGLEPVSTYSVSVSAFTKVGNGNQYSNVVEFTTQESVP
EAVRNIECVARDWQSVSVRWDPPRKTNGIIIHYMITVGGNSTKVSPRDPTYTFTKLLPN
TSYVFEVRASTSAGEGNESRCDISTLPETVPSAPTNVAFSNVQSTSATLTWTKPDTIFG
YFQNYKITTQLRAQKCREWEPEECIEHQKDQYLYEANQTEETVHGLKKFRWYRFQVAAS
TNVGYSNASEWISTQTLPGPPDGPPENVHVVATSPFGINISWSEPAVITGPTFYLIDVK
SVDDDDFNISFLKSNEENKTTEINNLEVFTRYSVVITAFVGNVSRAYTDGKSSAEVIIT
TLESVPKDPPNNMTFQKIPDEVTKFQLTFLPPSQPNGNIRVYQALVYREDDPTAVQIHN
FSIIQKTDTSIIAMLEGLKGGHTYNISVYAINSAGAGPKVQMRITMDIKAPARPKSKPI
PIRDATGKLLVTSTTITIRMPICYYNDDHGPIRNVQVLVAETGAQQDGNVTKWYDAYFN
KARPYFTNEGFPNPPCIEGKTKFSGNEEIYVIGADNACMIPGNEEKICNGPLKPKKQYL
FKFRATNVMGQFTDSEYSDPIKTLGEGLSERTVEIILSVTLCILSIILLGTAIFAFVRI
RQKQKEGGTYSPRDAEIIDTKFKLDQLITVADLELKDERLTRLLSYRKSIKPISKKSFL
QHVEELCTNSNLKFQEEFSELPKFLQDLSSTDADLPWNRAKNRFPNIKPYNNNRVKLIA
DVSLPGSDYINASYVSGYLCPNEFIATQGPLPGTVGDFWRMVWETRTKTLVMLTQCFEK
GRIRCHQYWPEDNKPVTVFGDIVITKLMEDIQIDWTIRDLKIERHGDCMTVRQCNFTGW
PEHGVPENTTPLIHFVKLVRTSRAHDTTPMVVHCSAGVGRTGVFIALDHLTQHINNHDF
VDIYGLVAELRSERMCMVQNLAQYIFLHQCILDLLSNKGGHQPVCFVNYSTLQKMDSLD
AMEGDVELEWEETTM

Fig. 4

MDFLIIFLLLFIGTSETQVDVSNVVPGTRYDITISSISTTYTSPVTRIVTTNVTKPGPPV
FLAGERVGSAGILLSWNTPPNPNGRIISYIVKYKEVCPWMQTVYTQVRSKPDSLEVLLTN
LNPGTTYEIKVAAENSAGIGVFSDPFLFQTAESAPGKVVNLTVEAYNASAVKLIWYLPRQ
PNGKITSFKISVKHARSGIVVKDVSIRVEDILTGKLPESNENSESFLWSTASPSPTLGRV
TPPSRTTHSSSTLTQNEISSVWKEPISFVVTHLRPYTTYLFEVSAATTEAGYIDSTIVRT
PESATAAGTQANRVWSRPLANSNRPAAEGP

Fig. 5

LLGGKLTNRKDIHTKNPSVHHHHQRPKVDKTTKMGKKQSRKTGNSKKQSTSPPPKDGRAV
FDLQLAEVESTQVRITWKKPRQPNGIINQYRVKVLVPETGIILENTLLTGNNEYINDPMA
PEIVNIVEPMVGLYEGSAEMSSDLHSLATFIYNSHPDKNFPARNRAEDQTSPVVTTRNQY
ITDIAAEQLSYVIRRLVPFTEHMISVSAFTIMGEGPPTVLSVRTRQQVPSSIKIINYKNI
SSSSILLYWDPPEYPNGKITHYTIYAMELDTNRAFQITTIDNSFLITGLKKYTKYKMRVA
ASTHVGESSLSEENDIFVRTSEDEPESSPQDVEVIDVTADEIRLKWSPPEKPNGIIIAYE
VLYKNIDTLYMKNTSTTDIILRNLRPHTLYNISVRSYTRFGHGNQVSSLLSVRTSETVPD
SAPENITYKNISSGEIELSFLPPSSPNGIIQKYTIYLKRSNGNEERTINTTSLTQNIKVL
KKYTQYIIEVSASTLKGEGVRSAPISILTEEDAPDSPPQDFSVKQLSGVTVKLSWQPPLE
PNGIILYYTVYVWNRSSLKTINVTETSLELSDLDYNVEYSAYVTASTRFGDGKTRSNIIS
FQTPEGAPSDPPKDVYYANLSSSSIILFWTPPSKPNGIIQYYSVYYRNTSGTFMQNFTLH
EVTNDFDNMTVSTIIDKLTIFSYYTFWLTASTSVGNGNKSSDIIEVYTDQDIPEGFVGNL
TYESISSTAINVSWVPPAQPNGLVFYYVSLILQQTPRHVRPPLVTYERSIYFDNLEKYTD
YILKITPSTEKGFSDTYTAQLYIKTEEDVPETSPIINTFKNLSSTSVLLSWDPPVKPNGA
IISYDLTLQGPNENYSFITSDNYIILEELSPFTLYSFFAAARTRKGLGPSSILFFYTDES
VPLAPPQNLTLINCTSDFVWLKWSPSPLPGGIVKVYSFKIHEHETDTIYYKNISGFKTEA
KLVGLEPVSTYSIRVSAFTKVGNGNQFSNVVKFTTQESVPDVVQNMQCMATSWQSVLVKW
DPPKKANGIITQYMVTVERNSTKVSPQDHMYTFIKLLANTSYVFKVRASTSAGEGDESTC
HVSTLPETVPSVPTNIAFSDVQSTSATLTWIRPDTILGYFQNYKITTQLRAQKCKEWESE
ECVEYQKIQYLYEAHLTEETVYGLKKFRWYRFQVAASTNAGYGNASNWISTKTLPGPPDG
PPENVHVVATSPFSISISWSEPAVITGPTCYLIDVKSVDNDEFNISFIKSNEENKTIEIK
DLEIFTRYSVVITAFTGNISAAYVEGKSSAEMIVTTLESAPKDPPNNMTFQKIPDEVTKF
QLTFLPPSQPNGNIQVYQALVYREDDPTAVQIHNLSIIQKTNTFVIAMLEGLKGGHTYNI
SVYAVNSAGAGPKVPMRITMDIKAPARPKTKPTPIYDATGKLLVTSTTITIRMPICYYSD
DHGPIKNVQVLVTETGAQHDGNVTKWYDAYFNKARPYFTNEGFPNPPCTEGKTKFSGNEE
IYIIGADNACMIPGNEDKICNGPLKPKKQYLFKFRATNIMGQFTDSDYSDPVKTLGEGLS
ERTVEIILSVTLCILSIILLGTAIFAFARIRQKQKEGGTYSPQDAEIIDTKLKLDQLITV
ADLELKDERLTRLLSYRKSIKPISKKSFLQHVEELCTNNNLKFQEEFSELPKFLQDLSST
DADLPWNRAKNRFPNIKPYNNNRVKLIADASVPGSDYINASYISGYLCPNEFIATQGPLP
GTVGDFWKMVWETRAKTLVMLTQCFEKGRIRCHQYWPEDNKPVTVFGDIVITKLMEDVQI
DWTIRDLKIERHGDCMTVRQCNFTAWPEHGVPENSAPLIHFVKLVRASRAHDTTPMIVHC
SAGVGRTGVFIALDHLTQHINDHDFVDIYGLVAELRSERMCMVQNLAQYIFLHQCILDLL
SNKGSNQPICFVNYSALQKMDSLDAMEGDVELEWEETTM

Fig. 6

ATGGATTTTCTTATCATTTTTCTTTTACTTTTTATTGGGACTTCAGAGACACAGGTTGAT
GTTTCCAATGTCGTTCCTGGTACTAGGTACGATATAACCATCTCTTCAATTTCTACAACA
TACACCTCACCTGTTACTAGAATAGTGACAACAAATGTAACAAAACCAGGGCCTCCAGTC
TTCCTAGCCGGGGAAAGAGTCGGATCTGCTGGGATTCTTCTGTCTTGGAATACACCACCT
AATCCAAATGGAAGGATTATATCTTACATTGTCAAATATAAGGAAGTTTGTCCGTGGATG
CAAACAGTATATACACAAGTCAGATCAAAGCCAGACAGTCTGGAAGTTCTTCTTACTAAT
CTTAATCCTGGAACAACATATGAAATTAAGGTTGCTGCTGAAAACAGTGCTGGCATTGGA
GTGTTTAGTGATCCATTTCTCTTCCAAACTGCAGAAAGTGCTCCAGGAAAAGTGGTGAAT
CTCACAGTTGAGGCCTACAACGCTTCAGCAGTTAAGCTGATTTGGTATTTACCTCGGCAA
CCAAATGGCAAAATTACCAGCTTCAAGATTAGTGTCAAGCATGCCAGAAGTGGGATAGTA
GTGAAAGATGTCTCAATCAGAGTAGAGGACATTTTGACTGGGAAATTGCCAGAAAGCAAT
GAGAATAGTGAATCTTTTTTATGGAGTACAGCCAGCCCTTCTCCAACCCTTGGTAGAGTT
ACACCTCCATCGCGTACCACACATTCATCAAGCACGTTGACACAGAATGAGATCAGCTCT
GTGTGGAAAGAGCCTATCAGTTTTGTAGTGACACACTTGAGACCTTATACAACATATCTT
TTTGAAGTTTCAGCTGCTACAACTGAAGCAGGTTATATTGATAGTACGATTGTCAGAACA
CCAGAATCAGCCACCGCTGCTGGTACCCAGGCTAACAGGGTCTGGAGTAGACCTCTAGCA
AACTCCAACAGACCTGCAGCGGAGGGTCCT

Fig. 7

CTGTTAGGAGGAAAACTAACAAACAGAAAGGACATCCACACCAAAAACCCATCTGTACAT
CACCATCATCAAAGACCAAAAGTAGATAAAACCACAAAGATGGGGAAAAAACAGAGCAGA
AAAACTGGAAACTCTAAAAAGCAGAGCACCTCTCCTCCTCCAAAGGATGGCAGGGCAGTG
TTTGATTTACAACTTGCAGAGGTAGAATCCACGCAAGTAAGAATTACTTGGAAGAAACCA
CGACAACCAAATGGAATTATTAACCAATACCGAGTGAAAGTGCTAGTTCCAGAGACAGGA
ATAATTTTGGAAAATACTTTGCTCACTGGAAATAATGAGTATATAAATGACCCCATGGCT
CCAGAAATTGTGAACATAGTAGAGCCAATGGTAGGATTATATGAGGGTTCAGCAGAGATG
TCGTCTGACCTTCACTCACTTGCTACATTTATATATAACAGCCATCCAGATAAAAACTTT
CCTGCAAGGAATAGAGCTGAAGACCAGACTTCACCAGTTGTAACTACAAGGAATCAGTAT
ATTACTGACATTGCAGCTGAACAGCTGTCTTATGTTATCAGGAGACTTGTACCTTTCACT
GAGCACATGATTAGTGTATCTGCTTTCACCATCATGGGAGAAGGACCACCAACAGTTCTC
AGTGTTAGGACACGTCAGCAAGTGCCAAGCTCCATTAAAATTATAAACTATAAAAATATT
AGTTCTTCATCTATTTTGTTATATTGGGATCCTCCAGAATATCCCAATGGAAAAATAACT
CACTATACGATTTATGCAATGGAATTGGATACAAACAGAGCATTCCAGATAACTACCATA
GATAACAGCTTTCTCATAACAGGGTTAAAGAAATACACAAAATACAAAATGAGAGTGGCA
GCCTCAACCCACGTTGGAGAAAGTTCTTTGTCTGAAGAAAATGACATCTTTGTGAGAACT
TCAGAAGATGAACCGGAATCATCACCTCAAGATGTCGAAGTAATTGATGTTACCGCAGAT
GAAATAAGGTTGAAGTGGTCACCACCCGAAAAGCCCAATGGGATCATTATTGCTTATGAA
GTGCTATATAAAAATATAGATACTTTATATATGAAGAACACATCAACAACAGACATAATA
TTAAGGAACTTAAGACCTCACACCCTCTATAACATTTCTGTAAGGTCTTACACCAGATTT
GGTCATGGCAATCAGGTATCTTCTTTACTCTCTGTAAGGACTTCGGAGACTGTGCCTGAT
AGTGCACCAGAAAATATCACTTACAAAAATATTTCTTCTGGAGAGATTGAGCTATCATTC
CTTCCCCCAAGTAGTCCCAATGGAATCATACAAAAATATACAATTTATCTCAAGAGAAGT
AATGGAAATGAGGAAAGAACTATAAATACAACCTCTTTAACCCAAAACATTAAAGTACTG
AAGAAATATACCCAATATATCATTGAGGTGTCTGCTAGTACACTCAAAGGTGAAGGAGTT
CGGAGTGCTCCCATAAGTATACTGACGGAGGAAGATGCTCCTGATTCTCCCCCTCAAGAC
TTCTCTGTAAAACAGTTGTCTGGTGTCACGGTGAAGTTGTCATGGCAACCACCCCTGGAG
CCAAATGGAATTATCCTTTATTACACAGTTTATGTCTGGAATAGATCATCATTAAAAACT
ATTAATGTCACTGAAACATCATTGGAGTTATCAGATTTGGATTATAATGTTGAATACAGT
GCTTATGTAACAGCTAGCACCAGATTTGGTGATGGGAAAACAAGAAGCAATATCATTAGC
TTTCAAACACCAGAGGGAGCACCAAGCGATCCTCCCAAAGATGTTTATTATGCAAACCTC
AGTTCTTCATCAATAATTCTTTTCTGGACACCTCCTTCAAAACCTAATGGGATTATACAA
TATTACTCTGTTTATTACAGAAATACTTCAGGTACTTTTATGCAGAATTTTACACTCCAT
GAAGTAACCAATGACTTTGACAATATGACTGTATCCACAATTATAGATAAACTGACAATA
TTCAGCTACTATACATTTTGGTTAACAGCAAGTACTTCAGTTGGAAATGGGAATAAAAGC
AGTGACATCATTGAAGTATACACAGATCAAGACATACCTGAAGGGTTTGTTGGAAACCTG
ACTTACGAATCCATTTCGTCAACTGCAATAAATGTAAGCTGGGTCCCACCGGCTCAACCA
AACGGTCTAGTCTTCTACTATGTTTCACTGATCTTACAGCAGACTCCTCGCCATGTGAGA
CCACCTCTTGTTACATATGAGAGAAGCATATATTTTGATAATCTGGAAAAATACACTGAT
TATATATTAAAAATTACTCCATCAACAGAAAAGGGATTCTCTGATACCTATACTGCCCAG
CTATACATCAAGACTGAAGAAGATGTCCCAGAAACTTCACCAATAATCAACACTTTTAAA
AACCTTTCCTCTACCTCAGTTCTCTTATCATGGGATCCCCCAGTAAAGCCAAATGGTGCA
ATAATAAGTTATGATTTAACTTTACAAGGACCAAATGAAAATTATTCTTTCATTACTTCT
GATAATTACATAATATTGGAAGAGCTTTCACCATTTACATTATATAGCTTTTTTGCTGCC
GCAAGAACTAGAAAAGGACTTGGTCCTTCCAGTATTCTTTTCTTTTACACAGATGAGTCA
GTGCCGTTAGCACCTCCACAAAATTTGACTTTAATCAACTGTACTTCAGACTTTGTATGG
CTGAAATGGAGCCCAAGTCCTCTTCCAGGTGGTATTGTTAAAGTATATAGTTTTAAAATT
CATGAACATGAAACTGACACTATATATTATAAGAATATATCAGGATTTAAAACTGAAGCC
AAACTTGTTGGACTGGAACCAGTCAGCACCTACTCTATCCGTGTATCTGCGTTCACCAAA
GTTGGAAATGGCAATCAATTTAGTAATGTAGTAAAATTCACAACCCAAGAATCAGTTCCA
GATGTCGTGCAGAATATGCAGTGCATGGCAACTAGCTGGCAGTCAGTTTTAGTGAAATGG
GATCCACCCAAAAAGGCAAATGGAATAATAACGCAGTATATGGTAACAGTTGAAAGGAAT
TCTACAAAAGTTTCTCCCCAAGATCACATGTACACTTTCATAAAGCTTCTTGCCAATACC
TCATATGTCTTTAAAGTAAGAGCTTCAACCTCAGCTGGTGAAGGTGATGAAAGCACATGC
CATGTCAGCACACTACCTGAAACAGTTCCCAGTGTTCCCACAAATATTGCTTTTTCTGAT
GTTCAGTCAACTAGTGCAACATTGACATGGATAAGACCTGACACTATCCTTGGCTACTTT
CAAAATTACAAAATTACCACTCAACTTCGTGCTCAAAAAATGCAAAGAATGGGAATCCGAA
GAATGTGTTGAATATCAAAAAAATTCAATACCTCTATGAAGCTCACTTAACTGAAGAGACA

Fig. 7 (continued)

GTATATGGATTAAAGAAATTTAGATGGTATAGATTCCAAGTGGCTGCCAGCACCAATGCT
GGCTATGGCAATGCTTCAAACTGGATTTCTACAAAAACTCTGCCTGGCCCTCCAGATGGT
CCTCCTGAAAATGTTCATGTAGTAGCAACATCACCTTTTAGCATCAGCATAAGCTGGAGT
GAACCTGCTGTCATTACTGGACCAACATGTTATCTGATTGATGTCAAATCGGTAGATAAT
GATGAATTTAATATATCCTTCATCAAGTCAAATGAAGAAAATAAAACCATAGAAATTAAA
GATTTAGAAATATTCACAAGGTATTCTGTAGTGATCACTGCATTTACTGGGAACATTAGT
GCTGCATATGTAGAAGGGAAGTCAAGTGCTGAAATGATTGTTACTACTTTAGAATCAGCC
CCAAAGGACCCACCTAACAACATGACATTTCAGAAGATACCAGATGAAGTTACAAAATTT
CAATTAACGTTCCTTCCTCCTTCTCAACCTAATGGAAATATCCAAGTATATCAAGCTCTG
GTTTACCGAGAAGATGATCCTACTGCTGTCCAGATTCACAACCTCAGTATTATACAGAAA
ACCAACACATTCGTCATTGCAATGCTAGAAGGACTAAAAGGTGGACATACATACAATATC
AGTGTTTACGCAGTCAATAGTGCTGGTGCAGGTCCAAAGGTTCCGATGAGAATAACCATG
GATATCAAAGCTCCAGCACGACCAAAAACCAAACCAACCCCTATTTATGATGCCACAGGA
AAACTGCTTGTGACTTCAACAACAATTACAATCAGAATGCCAATATGTTACTACAGTGAT
GATCATGGACCAATAAAAAATGTACAAGTGCTTGTGACAGAAACAGGAGCTCAGCATGAT
GGAAATGTAACAAAGTGGTATGATGCATATTTTAATAAAGCAAGGCCATATTTTACAAAT
GAAGGCTTTCCTAACCCTCCATGTACAGAAGGAAAGACAAAGTTTAGTGGCAATGAAGAA
ATCTACATCATAGGTGCTGATAATGCATGCATGATTCCTGGCAATGAAGACAAAATTTGC
AATGGACCACTGAAACCAAAAAAGCAATACTTATTTAAATTTAGAGCTACAAATATTATG
GGACAATTTACTGACTCTGATTATTCTGACCCTGTTAAGACTTTAGGGGAAGGACTTTCA
GAAAGAACCGTAGAGATCATTCTTTCCGTCACTTTGTGTATCCTTTCAATAATTCTCCTT
GGAACAGCTATTTTTGCATTTGCAAGAATTCGACAGAAGCAGAAAGAAGGTGGCACATAC
TCTCCTCAGGATGCAGAAATTATTGACACTAAATTGAAGCTGGATCAGCTCATCACAGTG
GCAGACCTGGAACTGAAGGACGAGAGATTAACGCGGTTACTTAGTTATAGAAAATCCATC
AAGCCAATAAGCAAGAAATCCTTCCTGCAACATGTTGAAGAGCTTTGCACAAACAACAAC
CTAAAGTTTCAAGAAGAATTTTCGGAATTACCAAAATTTCTTCAGGATCTTTCTTCAACT
GATGCTGATCTGCCTTGGAATAGAGCAAAAAACCGCTTCCCAAACATAAAACCATATAAT
AATAACAGAGTAAAGCTGATAGCTGACGCTAGTGTTCCAGGTTCGGATTATATTAATGCC
AGCTATATTTCTGGTTATTTATGTCCAAATGAATTTATTGCTACTCAAGGTCCACTACCA
GGAACAGTTGGAGATTTTTGGAAAATGGTGTGGGAAACCAGAGCAAAAACATTAGTAATG
CTAACACAGTGTTTTGAAAAAGGACGGATCAGATGCCATCAGTATTGGCCAGAGGACAAC
AAGCCAGTTACTGTCTTTGGAGATATAGTGATTACAAAGCTAATGGAGGATGTTCAAATA
GATTGGACTATCAGGGATCTGAAAATTGAAAGGCATGGGGATTGCATGACTGTTCGACAG
TGTAACTTTACTGCCTGGCCAGAGCATGGGGTTCCTGAGAACAGCGCCCCTCTAATTCAC
TTTGTGAAGTTGGTTCGAGCAAGCAGGGCACATGACACCACACCTATGATTGTTCACTGC
AGTGCTGGAGTTGGAAGAACTGGAGTTTTTATTGCTCTGGACCATTTAACACAACATATA
AATGACCATGATTTTGTGGATATATATGGACTAGTAGCTGAACTGAGAAGTGAAAGAATG
TGCATGGTGCAGAATCTGGCACAGTATATCTTTTTACACCAGTGCATTCTGGATCTCTTA
TCAAATAAGGGAAGTAATCAGCCCATCTGTTTTGTTAACTATTCAGCACTTCAGAAGATG
GACTCTTTGGACGCCATGGAAGGTGATGTTGAGCTTGAATGGGAAGAAACCACTATG

Fig. 8

```
tttactgctt aaaatctagg tgggtgtgag gccatgactg tgaacgttgt cagtcactga
ctgtattcag ctgtaaaggc ttgacaaggt gagaaaaacc tcccacactt ttaaaatgaa
tgaaatgttt gaacactaag tagtcataca caaaataatt accttaattt catatgttgt
tccaggatta agattagtaa gaagaacttc cagactgtct ggctttgatc tgacttgtgt
atatactgtt tgcatccacg gacaaacttc cttatatttg acaatgtaag atataatcct
tccatttgga ttaggtggtg tattccaaga cagaagaatc ccagcagatc cgactctttc
cccggctagg aagactggag gccctggttc tatttttaaa aaatccattt tg
```

Fig. 9

```
Q:       1 MDFLIIFLLLFIGXSEXQVDVSNVVPGXRYDIXISSISXX-YXSPVXRIVXXNVXKPGPP
             MDF . FL.L.IG.SE.QVDVS: ..G. YDI.:SS:S.. Y.SPV.R.:..NV.KPGPP
H:       2 MDFHFSFLFLLIGTSESQVDVSSSFDGTGYDITLSSVSATTYSSPVSRTLATNVTKPGPP

           VFLAGERVGSAGILLSWNTPPNPNGRIISYIVKYKEVCPWMQTVYTQVRSKPDSLEVLLT
           VFLAGERVGSAGILLSWNTPPNPNGRIISY:VKYKEVCPWMQT.YT:.R:KPDSLEVLLT
           VFLAGERVGSAGILLSWNTPPNPNGRIISYVVKYKEVCPWMQTAYTRARAKPDSLEVLLT

           NLNPGTTYEIKVAAENSAGIGVFSDPFLFQTAESAPGKVVNLTVEAYNASAVKLIWYLPR
           NLNPGTTYEIKVAAEN:AGIGVFSDPFLFQTAESAPGKVVNLTVEA.N SAV.LIWYLPR
           NLNPGTTYEIKVAAENNAGIGVFSDPFLFQTAESAPGKVVNLTVEALNYSAVNLIWYLPR

           QPNGKITSFKISVKHARSGIVVKDVSIRVEDILTGKLPESNENSESFLWSTASPSPTLGR
           QPNGKITSFKISVKHARSGIVVKDVS:RVEDIL:GKLPE.NENSESFLWST.SPSPTLGR
           QPNGKITSFKISVKHARSGIVVKDVSLRVEDILSGKLPECNENSESFLWSTTSPSPTLGR

           VTPPSRTTHSSSTLTQNEISSVWKEPISFVVTHLRPYTTYLFEVSAATTEAGYIDSTIVR
           VTP. RTT.SSST..:::ISSVWKEPISFVVTHLRPYTTYLFEVSA.TTEAGYIDSTIVR
           VTPTVRTTQSSSTAARSKISSVWKEPISFVVTHLRPYTTYLFEVSAVTTEAGYIDSTIVR

           TPESVPEGPPQNCVTGNITGKSFSILWDPPTIVTGKFSYRVELYGPSGRILDNSTKDLKF
           TPESVPEGPPQNC:.GN:TGK:FSI WDPPTIVTGKFSYRVELYGPSGRILDNSTKDL:F
           TPESVPEGPPQNCIMGNVTGKAFSISWDPPTIVTGKFSYRVELYGPSGRILDNSTKDLRF

           AFTNLTPFTMYDVYIAAETSAGTGPKSNISVFTPPDVPGAVFDLQLAEVESTQVRITWKK
           AFT:LTPFTMYDVY:AAETSAG.GPKSN:SVFTPPDVPGAVFDLQ:AEVE:T::RITW:K
           AFTHLTPFTMYDVYVAAETSAGVGPKSNLSVFTPPDVPGAVFDLQIAEVEATEIRITWRK

           PRQPNGIINQYRVKVLVPETGIILENTLLTGNNEYINDPMAPEIVNIVEPMVGLYEGSAE
           PRQPNGII:QYRVKV V ETG::LENTLLTG.:E I::PM:PEI:N:V:PM:G.YEGS.E
           PRQPNGIISQYRVKVSVLETGVVLENTLLTGQDESISNPMSPEIMNLVDPMIGFYEGSGE

           MSSDLHSLATFIYNSHPDKNFPARNRAEDQTSPVVTTRNQYITDIAAEQLSYVIRRLVPF
           MSSDLHS A:FIYNSHP..:FPA..RAE:Q:SPVVTTRNQY:TDI.AEQLSYV:RRLVPF
           MSSDLHSPASFIYNSHPHNDFPASTRAEEQSSPVVTTRNQYMTDITAEQLSYVVRRLVPF

           TEHMISVSAFTIMGEGPPTVLSVRTRQQVPSSIKIINYKNISSSSILLYWDPPEYPNGKI
           TEH.ISVSAFTIMGEGPPTVL:VRTR:QVPSSI:IINYKNISSSSILLYWDPPEYPNGKI
           TEHTISVSAFTIMGEGPPTVLTVRTREQVPSSIQIINYKNISSSSILLYWDPPEYPNGKI

           THYTIYAMELDTNRAFQITTIDNSFLITGLKKYTKYKMRVAASTHVGESSLSEENDIFVR
           THYTIYA.ELDTNRAFQ:TT:DNSFLITGLKKYT:YKMRVAASTHVGESSLSEENDIFVR
           THYTIYATELDTNRAFQMTTVDNSFLITGLKKYTRYKMRVAASTHVGESSLSEENDIFVR

           TSEDEPESSPQDVEVIDVTADEIRLKWSPPEKPNGIIIAYEVLYKNIDTLYMKNTSTTDI
           T.EDEPESSPQDV:V..V:..E:RLKWSPPEKPNGIIIAYEVLY:N.DTL::KNTSTTDI
           TPEDEPESSPQDVQVTGVSPSELRLKWSPPEKPNGIIIAYEVLYQNADTLFVKNTSTTDI
```

Fig. 9 (continued).

```
ILRNLRPHTLYNISVRSYTRFGHGNQVSSLLSVRTSETVPDSAPENITYKNISSGEIELS
I:.:L:P:TLYNIS:RSYTR.GHGNQ SSLLSVRTSETVPDSAPENITYKNISSGEIE:S
IISDLKPYTLYNISIRSYTRLGHGNQSSSLLSVRTSETVPDSAPENITYKNISSGEIEIS

FLPPSSPNGIIQKYTIYLKRSNGNEERTINTTSLTQNIKVLKKYTQYIIEVSASTLKGEG
FLPP.SPNGIIQKYTIYLKRSN.:E.RTINTTSLTQ.I  LKKYT.Y:IEVSASTLKGEG
FLPPRSPNGIIQKYTIYLKRSNSHEARTINTTSLTQTIGGLKKYTHYVIEVSASTLKGEG
VRSAPISILTEEDAPDSPPQDFSVKQLSGVTVKLSWQPPLEPNGIILYYTVYVWNRSSLK
:RS.PISILTEEDAPDSPPQ:FSVKQLSGVTV.LSWQPPLEPNGIILYYTVYVW::SSL:
IRSRPISILTEEDAPDSPPQNFSVKQLSGVTVMLSWQPPLEPNGIILYYTVYVWDKSSLR

TINVTETSLELSDLDYNVEYSAYVTASTRFGDGKTRSNIISFQTPEGAPSDPPKDVYYAN
.IN.TE.SL LSDLDYNV:Y.A VTASTRFGDG..RS:II:F:TPEG.PSDPP.DV:Y.N
AINATEASLVLSDLDYNVDYGACVTASTRFGDGNARSSIINFRTPEGEPSDPPNDVHYVN

LSSSSIILFWTPPSKPNGIIQYYSVYYRNTSGTFMQNFTLHEVTNDFDNMTVSTIIDKLT
LSSSSIILFWTPP KPNGIIQYYSVYY:NTSGTF:QNFTL :VT.: DN:TVS. I :L.
LSSSSIILFWTPPVKPNGIIQYYSVYYQNTSGTFVQNFTLLQVTKESDNVTVSARIYRLA

IFSYYTFWLTASTSVGNGNKSSDIIEVYTDQDIPEGFVGNLTYESISSTAINVSWVPPAQ
IFSYYTFWLTASTSVGNGNKSSDII.VYTDQDIPEG VGNLT:ESISSTAI:VSW PP:Q
IFSYYTFWLTASTSVGNGNKSSDIIHVYTDQDIPEGPVGNLTFESISSTAIHVSWEPPSQ

PNGLVFYYVSLILQQTP-RHVRPPLVTYERSIYFDNLEKYTDYILKITPSTEKGFSDTYT
PNGLVFYY:SL LQQ:P RH: PPLVTYE.SI FD:LEKYTDYI.KITPSTEKGFS:TYT
PNGLVFYYLSLNLQQSPPRHMIPPLVTYENSIDFDDLEKYTDYIFKITPSTEKGFSETYT

AQLYIKTEEDVPETSPIINTFKNLSSTSVLLSWDPPVKPNGAIISYDLTLQGPNENYSFI
.QL:IKTEEDVP:T.PIINTFKNLSSTS:LLSWDPP:KPNGAI:.Y.LTLQGP:.N::F:
TQLHIKTEEDVPDTPPIINTFKNLSSTSILLSWDPPLKPNGAILGYHLTLQGPHANHTFV

TSDNYIILEELSPFTLYSFFAAARTRKGLGPSSILFFYTDESVPLAPPQNLTLINCTSDF
TS.N:I:LEELSPFTLYSFFAAART.KGLGPSSILFFYTDES.PLAPPQNLTLIN TSDF
TSGNHIVLEELSPFTLYSFFAAARTMKGLGPSSILFFYTDESAPLAPPQNLTLINYTSDF

VWLKWSPSPLPGGIVKVYSFKIHEHETDTIYYKNISGFKTEAKLVGLEPVSTYSIRVSAF
VWL.WSPSPLPGGIVKVYSFKIHEHETDT::YKNISG.:T:AKL GLEPVSTYS:.VSAF
VWLTWSPSPLPGGIVKVYSFKIHEHETDTVFYKNISGLQTDAKLEGLEPVSTYSVSVSAF

TKVGNGNQFSNVVKFTTQESVPDVVQNMQCMATSWQSVLVKWDPPKKANGIITQYMVTVE
TKVGNGNQ:SNVV:FTTQESVP:.V:N::C:A..WQSV V:WDPP:K.NGII..YM:TV
TKVGNGNQYSNVVEFTTQESVPEAVRNIECVARDWQSVSVRWDPPRKTNGIIIHYMITVG

RNSTKVSPQDHMYTFIKLLANTSYVFKVRASTSAGEGDESTCHVSTLPETVPSVPTNIAF
 NSTKVSP:D .YTF.KLL.NTSYVF:VRASTSAGEG:ES.C.:STLPETVPS.PTN:AF
GNSTKVSPRDPTYTFTKLLPNTSYVFEVRASTSAGEGNESRCDISTLPETVPSAPTNVAF

SDVQSTSATLTWIRPDTILGYFQNYKITTQLRAQKCKEWESEECVEYQKIQYLYEAHLTE
S:VQSTSATLTW.:PDTI.GYFQNYKITTQLRAQKC:EWE.EEC:E:QK QYLYEA: TE
SNVQSTSATLTWTKPDTIFGYFQNYKITTQLRAQKCREWEPEECIEHQKDQYLYEANQTE

ETVYGLKKFRWYRFQVAASTNAGYGNASNWISTKTLPGPPDGPPENVHVVATSPFSISIS
ETV:GLKKFRWYRFQVAASTN.GY.NAS.WIST:TLPGPPDGPPENVHVVATSPF.I:IS
ETVHGLKKFRWYRFQVAASTNVGYSNASEWISTQTLPGPPDGPPENVHVVATSPFGINIS
```

Fig. 9 (continued)

```
WSEPAVITGPTCYLIDVKSVDNDEFNISFIKSNEENKTIEIKDLEIFTRYSVVITAFTGN
WSEPAVITGPT YLIDVKSVD:D:FNISF:KSNEENKT.EI.:LE:FTRYSVVITAF.GN
WSEPAVITGPTFYLIDVKSVDDDDFNISFLKSNEENKTTEINNLEVFTRYSVVITAFVGN

<ATP/GTP-binding site motif>
ISAAYVEGKSSAEMIVTTLESAPKDPPNNMTFQKIPDEVTKFQLTFLPPSQPNGNIQVYQ
:S.AY.:GKSSAE:I:TTLES.PKDPPNNMTFQKIPDEVTKFQLTFLPPSQPNGNI:VYQ
VSRAYTDGKSSAEVIITTLESVPKDPPNNMTFQKIPDEVTKFQLTFLPPSQPNGNIRVYQ

ALVYREDDPTAVQIHNLSIIQKTNTFVIAMLEGLKGGHTYNISVYAVNSAGAGPKVPMRI
ALVYREDDPTAVQIHN.SIIQKT:T  :IAMLEGLKGGHTYNISVYA:NSAGAGPKV.MRI
ALVYREDDPTAVQIHNFSIIQKTDTSIIAMLEGLKGGHTYNISVYAINSAGAGPKVQMRI

TMDIKAPARPKTKPTPIYDATGKLLVTSTTITIRMPICYYSDDHGPIKNVQVLVTETGAQ
TMDIKAPARPK:KP.PI DATGKLLVTSTTITIRMPICYY:DDHGPI:NVQVLV.ETGAQ
TMDIKAPARPKSKPIPIRDATGKLLVTSTTITIRMPICYYNDDHGPIRNVQVLVAETGAQ

HDGNVTKWYDAYFNKARPYFTNEGFPNPPCTEGKTKFSGNEEIYIIGADNACMIPGNEDK
.DGNVTKWYDAYFNKARPYFTNEGFPNPPC.EGKTKFSGNEEIY:IGADNACMIPGNE:K
QDGNVTKWYDAYFNKARPYFTNEGFPNPPCIEGKTKFSGNEEIYVIGADNACMIPGNEEK

                                   < transmembrane region >
ICNGPLKPKKQYLFKFRATNIMGQFTDSDYSDPVKTLGEGLSERTVEIILSVTLCILSII
ICNGPLKPKKQYLFKFRATN:MGQFTDS:YSDP:KTLGEGLSERTVEIILSVTLCILSII
ICNGPLKPKKQYLFKFRATNVMGQFTDSEYSDPIKTLGEGLSERTVEIILSVTLCILSII

LLGTAIFAFARIRQKQKEGGTYSPQDAEIIDTKLKLDQLITVADLELKDERLTRLLSYRK
LLGTAIFAF.RIRQKQKEGGTYSP:DAEIIDTK.KLDQLITVADLELKDERLTRLLSYRK
LLGTAIFAFVRIRQKQKEGGTYSPRDAEIIDTKFKLDQLITVADLELKDERLTRLLSYRK

SIKPISKKSFLQHVEELCTNNNLKFQEEFSELPKFLQDLSSTDADLPWNRAKNRFPNIKP
SIKPISKKSFLQHVEELCTN:NLKFQEEFSELPKFLQDLSSTDADLPWNRAKNRFPNIKP
SIKPISKKSFLQHVEELCTNSNLKFQEEFSELPKFLQDLSSTDADLPWNRAKNRFPNIKP

YNNNRVKLIADASVPGSDYINASYISGYLCPNEFIATQGPLPGTVGDFWKMVWETRAKTL
YNNNRVKLIAD.S:PGSDYINASY:SGYLCPNEFIATQGPLPGTVGDFW:MVWETR.KTL
YNNNRVKLIADVSLPGSDYINASYVSGYLCPNEFIATQGPLPGTVGDFWRMVWETRTKTL

VMLTQCFEKGRIRCHQYWPEDNKPVTVFGDIVITKLMEDVQIDWTIRDLKIERHGDCMTV
VMLTQCFEKGRIRCHQYWPEDNKPVTVFGDIVITKLMED:QIDWTIRDLKIERHGDCMTV
VMLTQCFEKGRIRCHQYWPEDNKPVTVFGDIVITKLMEDIQIDWTIRDLKIERHGDCMTV

<Tyrosine phosphatase signature: C is the active site residue>
RQCNFTAWPEHGVPENSAPLIHFVKLVRASRAHDTTPMIVHCSAGVGRTGVFIALDHLTQ
RQCNFT.WPEHGVPEN:.PLIHFVKLVR.SRAHDTTPM:VHCSAGVGRTGVFIALDHLTQ
RQCNFTGWPEHGVPENTTPLIHFVKLVRTSRAHDTTPMVVHCSAGVGRTGVFIALDHLTQ

HINDHDFVDIYGLVAELRSERMCMVQNLAQYIFLHQCILDLLSNKGSNQPICFVNYSALQ
HIN:HDFVDIYGLVAELRSERMCMVQNLAQYIFLHQCILDLLSNKG.:QP:CFVNYS.LQ
HINNHDFVDIYGLVAELRSERMCMVQNLAQYIFLHQCILDLLSNKGGHQPVCFVNYSTLQ

KMDSLDAMEGDVELEWEETTM    2299
KMDSLDAMEGDVELEWEETTM
KMDSLDAMEGDVELEWEETTM    2302
```

Fig. 10

```
Q:   1991 HVEELCTNNNLKFQEEFSELPKFLQDLSSTDADLPWNRAKNRFPNIKPYNNNRVKLIADA
          ::E.L .N:.LKF.:E:..:..  Q..: .:::L. N:.KNR:.N: .Y:::RV L.:
H:      8 NIERLKANDGLKFSQEYESIDPG-QQFTWENSNLEVNKPKNRYANVIAYDHSRVILTSID

          SVPGSDYINASYISGYLCPNEFIATQGPLPGTVGDFWKMVWETRAKTLVMLTQCFEKGRI
          .VPGSDYINA:YI.GY  .N.:IATQGPLP T:GDFW:MVWE.R..T:VM:T:. EK.R:
          GVPGSDYINANYIDGYRKQNAYIATQGPLPETMGDFWRMVWEQRTATVVMMTRLEEKSRV

          RCHQYWPEDNKPVTVFGDIVITKLMEDVQI-DWTIRDLKIERHG--DCMTVRQCNFTAWP
          :C.QYWP.  : ... G I :T L::.V:: .:T:R...:.:.G  : ..:RQ .F.AWP
          KCDQYWPA--RGTETCGLIQVT-LLDTVELATYTVRTFALHKSGSSEKRELRQFQFMAWP

          EHGVPENSAPLIHFVKLVRASRAHDTTPMIVHCSAGVGRTGVFIALDHLTQHINDHDFVD
          :HGVPE ..P:: F:: V:A... D. PM:VHCSAGVGRTG.FI.:D :.:.:.... VD
          DHGVPEYPTPILAFLRRVKACNPLDAGPMVVHCSAGVGRTGCFIVIDAMLERMKHEKTVD

          IYGLVAELRSERMCMVQNLAQYIFLHQCILDLLSNKGSNQPICFVNYSALQKMDSLDAME
          IYG V. :RS:R  MVQ.  QY:F:H:.:L:..:    :. P.  : Y: :QK:..:.. E
          IYGHVTCMRSQRNYMVQTEDQYVFIHEALLEAATCGHTEVPARNL-YAHIQKLGQVPPGE

          GDVELEWE    2295
          . ..:E E
          SVTAMELE     310
```

Fig. 11

```
Q:    2022 ADLPWNRAKNRFPNIKPYNNNRVKLIADASVPGSDYINASYISGYLCPNEFIATQGPLPG
           A:LP N: KNR. NI.PY.  .RV L .. ..V.GSDYINAS::.GY   ...:IATQGPL.
H:     323 ANLPCNKFKNRLVNIMPYELTRVCLQPIRGVEGSDYINASFLDGYRQQKAYIATQGPLAE

           TVGDFWKMVWETRAKTLVMLTQCFEKGRIRCHQYWPEDNKPVTVFGDIVITKLMEDVQID
           :. DFW:M:WE .:..:VMLT:. E.GR :CHQYWP.:..   . :   .V:..:.E  . .
           STEDFWRMLWEHNSTIIVMLTKLREMGREKCHQYWPAERS--ARYQYFVVDPMAEYNMPQ

           WTIRDLKI--ERHGDCMTVRQCNFTAWPEHGVPENSAPLIHFVKLVRASRAH--DTTPMI
           :.:R:.K:  .R.G...T:RQ .FT WPE.GVP:... .I.F:  V..::..  .. P:.
           YILREFKVTDARDGQSRTIRQFQFTDWPEQGVPKTGEGFIDFIGQVHKTKEQFGQDGPIT

           VHCSAGVGRTGVFIALDHLTQHINDHDFVDIYGLVAELRSERMCMVQNLAQYIFLHQCIL
           VHCSAGVGRTGVFI.L. :.:.:. ...VD:: .V..LR::R .MVQ.  QY ..::..L
           VHCSAGVGRTGVFITLSIVLERMRYEGVVDMFQTVKTLRTQRPAMVQTEDQYQLCYRAAL

           DLLSN    2262
           :.L.:
           EYLGS     565
```

Fig. 12

```
Q:   2027 NRAKNRFPNIKPYNNNRVKLIadasvPGSDYINASYISG--------YLCPNEFIATQGP
          N:.KNR:.:I PY::.RVKL      .GSDYINA:YI.G        Y  P..:IATQGP
H:      1 NkeKNRykdilpyDhtRVkLk.....egsDYINAnyidgpqsrtlvvykkpkayIaTQgP

          LPGTVGDFWKMVWETRAKTLVMLTQCFEKGR--IR-CHQYWPE---DNKPVTVFGDIVIT
          LP.T: DFW:MVWE.:...:VMLT:..EKGR   : C QYWPE   :.   ..:GDI.:T
          lpnTieDFWrMvweqkvtvIVMLtklvEkgreeeKtCaqYWPeeggeeggsltyGditVt

          KLMEDVQIDWTIRDLKIERhGDC--------------MTVRQCNFTAWPEHGVPEN-SAP
           :.E: . D:T:R.L::.. GD               .TV.. ::T.WP:HGVPE: ...
          lvseekdddytvRtleltn.gdeskneakqeklkgetrtVthyhytgWPDhgvPesgpks

          LIHFVKLVRASRA---------HDTT-------PMIVHCSAGVGRTGVFIALDHLTQHIN
          L:.F:: VR.S::          D:        P::VHCSAG:GRTG.FIA:D L.Q.:.
          lldflrkvrksqsqenesgaaspdsgprSsctaPivVHCSAGiGRTGtfiaidillqqle

          DHDFVDIYGLVAELRSERMCMVQNLAQYIFLHQCILD    2258
           ...VDI:.:V.ELRS:R  MVQ. .QYIF::..IL:
          aegvvdifdivkelRsqRpgmVQteeQYiFiydaile     271
```

Fig. 13

ATP_GTP_A region, transmembrane segments, and fibronectin type III domain region prediction.

```
High scoring transmembrane segments:
From 1906 to 1928:  length= 23, score=69.00  **
    1906  IILSVTLCIL SIILLGTAIF AFA
    L:  5(21.7%);  A:  3(13.0%);  I:  6(26.1%);

Fibronectin type III domain region
 from 1641 to 1734. Source: [pfamsearch]
 from 1538 to 1620. Source: [pfamsearch]
 from 1430 to 1526. Source: [pfamsearch]
 from 1340 to 1418. Source: [pfamsearch]
 from 1243 to 1328. Source: [pfamsearch]
 from 1150 to 1231. Source: [pfamsearch]
 from 1053 to 1134. Source: [pfamsearch]
 from 948 to 1040. Source: [pfamsearch]
 from 854 to 935. Source: [pfamsearch]
 from 759 to 842. Source: [pfamsearch]
 from 665 to 746. Source: [pfamsearch]
 from 569 to 651. Source: [pfamsearch]
 from 397 to 558. Source: [pfamsearch]
 from 305 to 386. Source: [pfamsearch]
 from 155 to 291. Source: [pfamsearch]
 from 56 to 143. Source: [pfamsearch]
```

Fig. 14

```
AF063249_1             2 MDFHFSFLFLLIGTSESQ                        VDVSSSFD
                         MDF + FL+L IGTSE+Q                        VDVS+
                         MDFLIIFLLLFIGTSETQ                        VDVSNVVP
chr12:88900001-32923 agtcaatctctagatgacGTATTTC  Intron 1   TAGgggtaggc
                         tattttttttttgccaca<0-----[32977:33317]-0>tatcattc
                         gtttctttatttgtagag                        tttctctt

AF063249_1            28 GTGYDITLSSVSATTYSSPVSRTLATNVT
                         GT YDIT+SS+S TTY+SPV+R + TNVT
                         GTRYDITISSIS-TTYTSPVTRIVTTNVT
chr12:88900001-33342 gaatgaaattat aatatcgaaagaaaga
                         gcgaatctcctc ccaccctcgttccatc
                         ttgctacctatt aaccatttaagaataa

AF063249_1            57                              PGPPVFLAGERVGSAGILLSWNT
                                                      PGPPVFLAGERVGSAGILLSWNT
                                     K:K[aaa]         PGPPVFLAGERVGSAGILLSWNT
chr12:88900001-33426 AGTGAGTA  Intron 2   TAGAAcgccgtcgggaggtggaccttaa
                          <1-----[33427:34067]-1>  cgcctttcgagtgccgtttcgac
                                                      agtaccacgaacattgttgtgta

AF063249_1            81 PPNPNGRIISYVVKYKEVCPWMQTAYTRARAKPDSLEVLLTNLNPGTTY
                         PPNPNGRIISY+VKYKEVCPWMQT YT+ R+KPDSLEVLLTNLNPGTTY
                         PPNPNGRIISYIVKYKEVCPWMQTVYTQVRSKPDSLEVLLTNLNPGTTY
chr12:88900001-34139 ccacagaaattagataggtctacagtacgatacgacggccaacacgaat
                         ccacaggttcattaaaatgcgtactacatgcacagtatttcatacgcca
                         attatagtatctcatgattgggaaataacaagactgatttttttaaat

AF063249_1           130 EIK                          VAAENNAGIGVFSDPFLFQTAES
                         EIK                          VAAEN+AGIGVFSDPFLFQTAES
                         EIK                          VAAENSAGIGVFSDPFLFQTAES
chr12:88900001-34286 gaaGTAATTA  Intron 3   TAGggggaaggaggtagctctcagga
                         ata<0-----[34295:65280]-0>tccaagcgtgttgactttaccag
                         atg                          tttacttctagtttatccataat

AF063249_1           156                              PGKVVNLTVEALNYSAVNLIWYL
                                                      PGKVVNLTVEA N SAV LIWYL
                                    A:A[gct]          PGKVVNLTVEAYNASAVKLIWYL
chr12:88900001-65350 GGTAATTT  Intron 4   CAGCTcgaggacagggtagtggacattt
                          <1-----[65351:66989]-1>  cgattatctacaaccctattgat
                                                      aaaggtcatgccctaatggtgta

AF063249_1           180 PRQPNGKITSFKISVKHARSGIVVKDVSLRVEDILSGKLPECN
                         PRQPNGKITSFKISVKHARSGIVVKDVS+RVEDIL+GKLPE N
                         PRQPNGKITSFKISVKHARSGIVVKDVSIRVEDILTGKLPESN
chr12:88900001-67061 ccccagaaaataaagacgaagaggaggtaagggatagatcgaa
                         cgacagatcgtatgtaacgggtttaatctgtaattcgatcaga
                         tgaatcatcccgttcgtcatgaagatcacaagctgtgagaact
```

Fig. 14 (continued)

```
AF063249_1        223                                ENSESFLWSTTSPSPTLGRVTPTVRT
                                                     ENSESFLWST SPSPTLGRVTP  RT
                                                     ENSESFLWSTASPSPTLGRVTPPSRT
chr12:88900001-67190 GTAAGTA  Intron 5   TAGgaagttttaagactcacgagacctca
                     <0-----[67190:68356]-0>aagacttggccgcccctggtccccgc
                                                     gttattagtaccttacttatatagtc

AF063249_1        249 TQSSSTAARSKISSVWKEPISFVVTHLRPYTTYLFEVSAVTTEAGYIDS
                      T SSST   +++ISSVWKEPISFVVTHLRPYTTYLFEVSA TTEAGYIDS
                      THSSSTLTQNEISSVWKEPISFVVTHLRPYTTYLFEVSAATTEAGYIDS
chr12:88900001-68435 acttaatacagaatgtagcaatggactactaatctggtggaagggtaga
                     caccgctcaaatgctgaactgtttcatgcaccattatcccccacgatag
                     ataacggagtgcctggagtcttagacgattaatttatattataattttt

AF063249_1        298 TIVRTPES                              VPEGPPQNCIMGNVT
                      TIVRTPES                                 G   N +
                      TIVRTPES           -:A[gcc]           TAAGTQANRVWSRPL
chr12:88900001-68582 aagaacgtGGTATGGT  Intron 6   CAGCCagggacgaagtaacc
                     cttgccac <1-----[68607:70433]-1>  cccgcacagtgggct
                     gtcaaaaa                               ctttcgtcgcgtata

AF063249_1        321 GKAFSISWDPPTIVTGKFSYRVELYGPSGRILDNSTKDLRFAFTHLTPF
                        +    + + P ++ GK++ R +++                TK+       H  P
                      ANSNRPAAEGP!LLGGKLTNRKDIH----------TKNPSVHHHHQRPK
chr12:88900001-70481 gataacggggc2ctggacaaaagac           aaactgcccccaca
                     cacagcccagc ttggatcagaata           caacctaaaaagca
                     acccataggtt gaaaaaacagccc           cacatatcttaaaa

AF063249_1        370 TMYDVYVAAETSAGVGPKSNLSVFTPPDVP
                          +   + S   G      S    PP
                      VDKTTKMGKKQSRKTGNSKKQSTSPPPKDG
chr12:88900001-70597 ggaaaaagaacaaaagataacaatcccagg
                     taaccatgaaaggacgacaaagcccccaag
                     atacagggaagcaatactaggcctttagtc

AF063249_1        400                                AVFDLQIAEVEATEIRITWRKPR
                                                     AVFDLQ+AEVE+T++RITW+KPR
                                 G:R[agg]            AVFDLQLAEVESTQVRITWKKPR
chr12:88900001-70687 AGTTCCTC  Intron 7   CAGGGggtgtccggggtacgaaataacc
                      <1-----[70688:72736]-1>  cttatatcataccatgtcgaacg
                                                     agttaatagaacgaaattggaaa

AF063249_1        424 QPNGIISQYRVKVSVLETGVVLENTLLTGQDE
                      QPNGII+QYRVKV V ETG++LENTLLTG +E
                      QPNGIINQYRVKVLVPETGIILENTLLTGNNE
chr12:88900001-72808 ccagaaactcgagcgcgagaatgaatcagaag
                     acagttaaagtatttcacgtttaacttcgaaa
                     aatattcacagagatagaaatgattgctattg
```

Fig. 14 (continued)

```
AF063249_1         456                           SISNPMSPEIMNLVDPMIGFYEGSGE
                                                  I++PM+PEI+N+V+PM+G+YEGS E
                                                 YINDPMAPEIVNIVEPMVGLYEGSAE
chr12:88900001-72904 GTATTGC  Intron 8    CAGtaagcagcgagaaggcaggttggtgg
                     <0-----[72904:84426]-0>ataactccattattacttgtaagcca
                                                 tatccgtaatgcaagagaaatgtaag

AF063249_1         482 MSSDLHSPASFIYNSHPHNDFPASTRAEEQSSPV
                       MSSDLHS A+FIYNSHP  +FPA  RAE+Q+SPV
                       MSSDLHSLATFIYNSHPDKNFPARNRAEDQTSPV
chr12:88900001-84505 attgcctcgatataaccgaatcgaaagggcatcg
                     tccatactccttaagacaaatccgagcaaaccct
                     ggtctcattatatcctatacttagtatacgtaat

AF063249_1         516                              TTRNQYMTDITAEQLSYVVRRLV
                                                    TTRNQY+TDI AEQLSYV+RRLV
                                V:V[gta]            TTRNQYITDIAAEQLSYVIRRLV
chr12:88900001-84607 GGTAGGTA  Intron 9    CAGTAaaaactaagagggccttgaaacg
                      <1-----[84608:86084]-1>   ccgaaatcatccaatcattggtt
                                                    tagtgtttctataggtttcgata

AF063249_1         540 PFTEHTISVSAFTIMGEGPPTVLTVRTREQ
                       PFTEH ISVSAFTIMGEGPPTVL+VRTR+Q
                       PFTEHMISVSAFTIMGEGPPTVLSVRTRQQ
chr12:88900001-86156 ctagcaaagtgtaaagggccagcagaaccc
                     ctcaattgtcctcttgagccccttgtgcgaa
                     tctgcgttattcccgaaaaaatcttgatga

AF063249_1         570                              PSSIQIINYKNISSSSILLYWDP
                                                    PSSI+IINYKNISSSSILLYWDP
                                V:V[gtg]            PSSIKIINYKNISSSSILLYWDP
chr12:88900001-86246 GGTAAGGA  Intron 10   AAGTGcataaaaataaaatttattttgc
                      <1-----[86247:86343]-1>   cgctattaaaatgccctttagac
                                                    acctatactattttattgatgtt

AF063249_1         594 PEYPNGKITHYTIYATELDTNRAFQMTTVDNSFLIT
                       PEYPNGKITHYTIYA ELDTNRAFQ+TT+DNSFLIT
                       PEYPNGKITHYTIYAMELDTNRAFQITTIDNSFLIT
chr12:88900001-86415 cgtcagaaactaatgagtgaaagtcaaaagaatcaa
                     caacagatcaactactatacagctatcctaagtttc
                     aatctaaatctgttagagtacaacgatcatcctcaa

AF063249_1         630                              LKKYTRYKMRVAASTHVGESSLS
                                                    LKKYT+YKMRVAASTHVGESSLS
                                G:G[ggg]            LKKYTKYKMRVAASTHVGESSLS
chr12:88900001-86523 GGTAGAAA  Intron 11   TAGGGtaataataaagggtacgggattt
                      <1-----[86524:87150]-1>   taaacaaatgtccccatgagctc
                                                    agacaacagagacacctaattgt
```

Fig. 14 (continued)

```
AF063249_1         654 EEENDIFVRTPED                                 PESSPQDVQVT
                       EEENDIFVRT ED                                 PESSPQDV+V
                       EEENDIFVRTSED            E:E[gaa]             PESSPQDVEVI
chr12:88900001-87222 ggagatgaatggGGTAAGAA  Intron 12   TAGAAcgttccgggga
                       aaaatttgccaa <1-----[87259:87401]-1>  cacccaatatt
                       aatcctgataat                                  gaaatatcaat

AF063249_1         678 GVSPSELRLKWSPPEKPNGIIIAYEVLYQNADTLFVKNTSTTDIIISDL
                        V+   E+RLKWSPPEKPNGIIIAYEVLY+N DTL++KNTSTTDII+ +L
                       DVTADEIRLKWSPPEKPNGIIIAYEVLYKNIDTLYMKNTSTTDIILRNL
chr12:88900001-87437 ggagggaatattccgacagaaagtggctaaagattaaaataagaataat
                       atccaatgtagcccaacagtttcaattaaatactataaccccatttgat
                       ttcataaggggaacagctgcttttagatatattatggcaaaacaaagca

AF063249_1         727 KPYTLYNISIRSYTRLGHGNQSSSLLSVRTSET
                       +P+TLYNIS+RSYTR GHGNQ SSLLSVRTSET
                       RPHTLYNISVRSYTRFGHGNQVSSLLSVRTSET
chr12:88900001-87584 accactaatgattaatgcgacgtttctgaatga
                       gcactaatctgcacgtgagaatccttctgccac
                       atccctcttagtccatttctgattactagtggt

AF063249_1         760                              PDSAPENITYKNISSGEIEISFL
                                                    PDSAPENITYKNISSGEIE+SFL
                                 V:V[gtg]           PDSAPENITYKNISSGEIELSFL
chr12:88900001-87683 GGTGAGCT  Intron 13   CAGTGcgagcgaaataaattggagcttc
                        <1-----[87684:97175]-1>  cagccaatcaaatccgatatctt
                                                    tttaaatctcattttagtgaact

AF063249_1         784 PPRSPNGIIQKYTIYLKRSNSHEARTINTTSLTQTIG
                       PP SPNGIIQKYTIYLKRSN +E RTINTTSLTQ I
                       PPSSPNGIIQKYTIYLKRSNGNEERTINTTSLTQNIK
chr12:88900001-97247 ccaacagaacataatcaaaagaggaaaaaattacaaa
                       ccggcagttaaactataggagaaagctaccctcaata
                       cattctacaaatattcgattatgaatatactacacta

AF063249_1         821                              LKKYTHYVIEVSASTLKGEGIRS
                                                    LKKYT Y+IEVSASTLKGEG+RS
                                 G:V[gta]           LKKYTQYIIEVSASTLKGEGVRS
chr12:88900001-97358 GGTAAAAG  Intron 14   TAGTAcaatactaaggtgaacaggggca
                        <1-----[97359:97720]-1>  taaacaattatccgctagagtgg
                                                    ggatcatctggtttacataatgt

AF063249_1         845 RPISILTEED                               PDSPPQNFSVKQL
                        PISILTEED                               PDSPPQ+FSVKQL
                       APISILTEED           A:A[gct]            PDSPPQDFSVKQL
chr12:88900001-97792 gcaaacagggGGTAAATA  Intron 15   CAGCTcgtcccgttgact
                       cctgttcaaa <1-----[97823:101475-1>   cacccaatctaat
                       tcatagggat                               tttctacctaagg
```

Fig. 14 (continued)

```
AF063249_1        869 SGVTVMLSWQPPLEPNGIILYYTVYV
                      SGVTV LSWQPPLEPNGIILYYTVYV
                      SGVTVKLSWQPPLEPNGIILYYTVYV           W:W[tgg]
chr12:88900001101517 tggagatttccccgcagaacttagtgTGGTAATAA  Intron 16
                      cgtctatcgacctacagtttaactat  <2-----[101597:125167
                      ttcggggagaacggatatcttcattc


AF063249_1        895     DKSSLRAINATEASLVLSDLDYNVDYGACVTASTRFGDGNARSSI
                          ++SSL+ IN TE SL LSDLDYNV+Y A VTASTRFGDG  RS+I
                          NRSSLKTINVTETSLELSDLDYNVEYSAYVTASTRFGDGKTRSNI
chr12:88900001125165 CAGGaatttaaaagagattgttgtgtaggtagtgagaaatgggaaaaaa
                     -2> agcctactatcacctatcataaataagcatccgcgtgagacggat
                         taaaaatttctaaaggaatgttttactttaatccatttgaaactc


AF063249_1        941 INFRTPEG                          PSDPPNDVHYVNLSS
                      I+F+TPEG                          PSDPP DV+Y NLSS
                      ISFQTPEG          E:A[gca]        PSDPPKDVYYANLSS
chr12:88900001125304 aatcacggGGTGAGTT   Intron 17   TAGCAcagccaggttgacat
                      tgtaccag <1-----[125329:12604-1>  cgaccaataacatgc
                      tctaaaga                          acttcattttacctt


AF063249_1        965 SSIILFWTPPVKPNGIIQYYSVYYQNTSGTFVQ
                      SSIILFWTPP KPNGIIQYYSVYY+NTSGTF+Q
                      SSIILFWTPPSKPNGIIQYYSVYYRNTSGTFMQ
chr12:88900001126092 ttaacttacctacagaactttgttaaatgatac
                      ccttttgccccacagttaaactaagaccgctta
                      aaattcgattaattgtaatctttcattatttgg


AF063249_1        998                           NFTLLQVTKESDNVTVSARIYRLAIF
                                                NFTL +VT + DN+TVS  I +L IF
                                                NFTLHEVTNDFDNMTVSTIIDKLTIF
chr12:88900001126191 GTAAGAA  Intron 18   CAGataccggaagtgaaagtaaagacaat
                      <0-----[126191:13092-0>atctaatcaataatctccttaatctt
                                                ttactaactctctgtacatatagaac


AF063249_1       1024 SYYTFWLTASTSVGNGNKSSDIIHVYTDQD
                      SYYTFWLTASTSVGNGNKSSDII VYTDQD
                      SYYTFWLTASTSVGNGNKSSDIIEVYTDQD
chr12:88900001131007 attatttagaatggagaaaagaaggtagcg
                      gaactgtccgcctgagaaggattatacaaa
                      cctatgaaattatatgtactcctaacatac


AF063249_1       1054                              PEGPVGNLTFESISSTAIHVSWE
                                                   PEG VGNLT+ESISSTAI+VSW
                                I:I[ata]           PEGFVGNLTYESISSTAINVSWV
chr12:88900001131097 AGTATGTA  Intron 19   TAGTAcggtggacatgtattagaagatg
                       <1-----[131098:13270-1>     cagttgatcaactcccctatggt
                                                   tagttacgtcactgatatacgc
```

Fig. 14 (continued)

```
AF063249_1        1078 PPSQPNGLVFYYLSLNLQQSPPRHMIPPLVTYENSIDFDDLEKYTDYIF
                       PP+QPNGLVFYY+SL LQQ+P RH+ PPLVTYE SI FD+LEKYTDYI+
                       PPAQPNGLVFYYVSLILQQTP-RHVRPPLVTYERSIYFDNLEKYTDYIL
chr12:88900001132779 ccgccagcgtttgtcatccac ccgacccgatgaaattgacgatagtat
                       cccacagtttaatctttaacc gatgccttcaaggtataataaacaatt
                       agtaactacccttagcaggtt ctgaatttatgacattttgaactttaa

AF063249_1        1127 KITPSTEKGFSETYTTQLHIKTEED
                       KITPSTEKGFS+TYT QL+IKTEED
                       KITPSTEKGFSDTYTAQLYIKTEED          V:V[gtc]
chr12:88900001132923 aaactagagttgatagcctaaagggGGTAGGCT  Intron 20
                       atccccaagtcacaccatatacaaa <1-----[132999:133241
                       attaaaagacttcttcgaccgtaat

AF063249_1        1152     PDTPPIINTFKNLSSTSILLSWDPPLKPNGAILGYHLTLQGPHA
                           P+T PIINTFKNLSSTS+LLSWDPP+KPNGAI+ Y LTLQGP+
                           PETSPIINTFKNLSSTSVLLSWDPPVKPNGAIISYDLTLQGPNE
chr12:88900001133239 TAGTCcgatcaaaataacttatgctttgccgacaggaaatgtatcgcag
                     -1>   caccccttactaatcccctttcgacctacagcttgaatctagcaa
                           aataaaccttactctcatcaagtcaagattaaatttataaaata

AF063249_1        1197 NHTFVTSGNHIVLEELSPFTLYSFFAAARTMKGLGPSSILFFYTDES
                       N++F+TS N+I+LEELSPFTLYSFFAAART KGLGPSSILFFYTDES
                       NYSFITSDNYIILEELSPFTLYSFFAAARTRKGLGPSSILFFYTDES
chr12:88900001133376 atttaatgataatggctctattattgggaaaagcgctaactttaggt
                       aacttccaaatttaatcctctagttcccgcgagtgccgttttacaac
                       tttctttttcaagagtaataatctttcaataaatttctttctcatga

AF063249_1        1244                            PLAPPQNLTLINYTSDFVWLTWS
                                                  PLAPPQNLTLIN TSDFVWL WS
                               A:V[gtg]           PLAPPQNLTLINCTSDFVWLKWS
chr12:88900001133517 GGTAAGCC  Intron 21  CAGTGctgcccatataatatgtgtcata
                      <1-----[133518:13388-1>     ctcccaatcttagccattgtagg
                                                  gaataatgtaccttactaggagc

AF063249_1        1268 PSPLPGGIVKVYSFKIHEHETDTVFYK
                       PSPLPGGIVKVYSFKIHEHETDT++YK
                       PSPLPGGIVKVYSFKIHEHETDTIYYK
chr12:88900001133954 caccccggagagtataacgcgagaattaGTAGGTT  Intron 22
                       cgctcggttatagtataaaacactaaa<0-----[134035:137708
                       atttattttaatttattatatctattg

AF063249_1        1295    NISGLQTDAKLEGLEPVSTYSVSVSAFTKVGNGNQYSNVVEFTTQE
                          NISG +T+AKL GLEPVSTYS+ VSAFTKVGNGNQ+SNVV+FTTQE
                          NISGFKTEAKLVGLEPVSTYSIRVSAFTKVGNGNQFSNVVKFTTQE
chr12:88900001137706 CAGaatgtaaggacggcgcgaattacgtgtaaggagactaaggataacg
                     -0>atcgtacacattgtactgcactgtcctcatgagaatgattatccaa
                          taaatatacattagaacccctctatgccatatctatttaaacacaa
```

Fig. 14 (continued)

```
AF063249_1      1341 S                          PEAVRNIECVARDWQSVSVRWD
                     S                          P+ V+N++C+A  WQSV V+WD
                     S          V:V[gtt]        PDVVQNMQCMATSWQSVLVKWD
chr12:88900001137847 tGGTTAGAT  Intron 23  AAGTTcgggcaactagaatctgtgatg
                     c <1-----[137851:14061-1>  cattaatagtccggactttaga
                     a                          atcggtggcgatcggatagagt

AF063249_1      1365 PPRKTNGIIIHYMITVGGNSTKVSPRDPTYTFTKLLPNTSYVFEVRAST
                     PP+K NGII  YM+TV  NSTKVSP+D  YTF KLL NTSYVF+VRAST
                     PPKKANGIITQYMVTVERNSTKVSPQDHMYTFIKLLANTSYVFKVRAST
chr12:88900001140686 ccaagagaaactagaggaataagtccgcatataaccgaattgtagagta
                     ccaacagttcaattctagaccatccaaatacttattcaccattatgccc
                     acagataaaggtgaatagttaattcatcgctcagttctcatctaaatac

AF063249_1      1414 SAGEGNESRCDISTLPET                          PSAPT
                     SAGEG+ES C +STLPET                          PS PT
                     SAGEGDESTCHVSTLPET          V:V[gtt]        PSVPT
chr12:88900001140833 tggggggaatcgaaccgaGGTAACTA  Intron 24  TAGTTcagca
                     ccgagaagcgatgctcac <1-----[140888:17928-1>  cgtcc
                     attattacactccaataa                          cttca

AF063249_1      1438 NVAFSNVQSTSATLTWTKPDTIFGYFQNYKITTQLRAQKCREWEPEECI
                     N+AFS+VQSTSATLTW +PDTI+GYFQNYKITTQLRAQKC+EWE EEC+
                     NIAFSDVQSTSATLTWIRPDTILGYFQNYKITTQLRAQKCKEWESEECV
chr12:88900001179299 aagttggctaagatataacgaacgttcataaaacccgcatagtgtggtg
                     atctcataccgcctcgtgcacttgataaaatccatgcaagaagacaagt
                     tttttttgattaagagaatctctcctatcatctatttaacaagacaatt

AF063249_1      1487 EHQKDQYLYEANQTEETVHGLKKFRWYRFQVAASTNVGYSNASEWISTQ
                     E+QK QYLYEA+ TEETV+GLKKFRWYRFQVAASTN GY NAS WIST+
                     EYQKIQYLYEAHLTEETVYGLKKFRWYRFQVAASTNAGYGNASNWISTK
chr12:88900001179446 gtcaactctggctaggagtgtaatattatcgggaaaggtgagtatataa
                     aaaataataacatcaactagtaatggagtatccgcacgagaccagtcca
                     ataatacctatcatagaataagatagtacagtcccttctcttacgttaa

AF063249_1      1536 TLPG                          PDGPPENVHVVATSPFGIN
                     TLPG                          PDGPPENVHVVATSPF I+
                     TLPG          P:P[cct]        PDGPPENVHVVATSPFSIS
chr12:88900001179593 accgCGTGAGTA  Intron 25  TAGCTcggccgagcgggatctaaa
                     ctcg <1-----[179606:19621-1>  cagccaatattcccctgtg
                     tgtc                          attttatttaaaaattccc

AF063249_1      1560 ISWSEPAVITGPTFYLIDVKS                       VDDDD
                     ISWSEPAVITGPT YLIDVKS                       VD+D+
                     ISWSEPAVITGPTCYLIDVKS                       VDNDE
chr12:88900001196271 aatagcggaagcattcaggatGTAAGGC  Intron 26  TAGggagg
                     tgggaccttcgccgattatac<0-----[196334:20159-0>taaaa
                     acgtattcttaaattgttcag                       attta
```

Fig. 14 (continued)

```
AF063249_1          1586 FNISFLKSNEENKTTEINNLEVFTRYSVVITAFVGNVSRAYTDGKSSAE
                         FNISF+KSNEENKT EI +LE+FTRYSVVITAF GN+S AY +GKSSAE
                         FNISFIKSNEENKTIEIKDLEIFTRYSVVITAFTGNISAAYVEGKSSAE
chr12:88900001201607 taattaataggaaaagaagtgataattggaagtagaaaggtgggatagg
                         tatcttacaaaaactataatattcgactttcctcgatgccatagacgca
                         ttacccgataatacaatataaacagttagctattgctttataaggatta

AF063249_1          1635 VIITTLES                              PKDPPNNMTFQKIPD
                         +I+TTLES                              PKDPPNNMTFQKIPD
                         MIVTTLES          V:A[gcc]            PKDPPNNMTFQKIPD
chr12:88900001201754 aagaatgtGGTAAGGA  Intron 27  TAGCCcagccaaaatcaacg
                         tttcctac <1-----[201779:20474-1>  caaccaatctaatca
                         gttttaaa                              agcatccgatggaat

AF063249_1          1659 EVTKFQLTFLPPSQPNGNIRVYQALVYREDDPTAVQIHNFSIIQKTDTS
                         EVTKFQLTFLPPSQPNGNI+VYQALVYREDDPTAVQIHN+SIIQKT+T
                         EVTKFQLTFLPPSQPNGNIQVYQALVYREDDPTAVQIHNLSIIQKTNTF
chr12:88900001204792 ggaatctatccctccagaacgtcgcgtcgggcaggcacacaaacaaaat
                         atcatatcttcccacagatataacttagaaaccctataatgttaacact
                         ataataagcttttattatcaatatgtcaatttttcgtcccttagaccac

AF063249_1          1708 IIAMLEGLKGGHTYNIS                           VYAINSAGA
                         +IAMLEGLKGGHTYNIS                           VYA+NSAGA
                         VIAMLEGLKGGHTYNIS                           VYAVNSAGA
chr12:88900001204939 gagacggcaggcataaaGTAAGAA  Intron 28  TAGgtggaaggg
                         ttcttagtaggacaatg<0-----[204990:20735-0>tactagcgc
                         ctagaaaaatatactct                           tcactttta

AF063249_1          1734 GPKVQMRITMDIK                          PARPKSKPIP
                         GPKV MRITMDIK                          PARPK+KP P
                         GPKVPMRITMDIK          A:A[gct]        PARPKTKPTP
chr12:88900001207380 gcagcaaaaagaaGGTACATA  Intron 29  TAGCTcgccaaacac
                         gcatctgtctata <1-----[207420:21053-1>  ccgcacaccc
                         tagtggaacgtca                          aaaaacaact

AF063249_1          1758 IRDATGKLLVTSTTITIRMPICYYNDDHGPIRNVQVLVAETG
                         I DATGKLLVTSTTITIRMPICYY+DDHGPI+NVQVLV ETG
                         IYDATGKLLVTSTTITIRMPICYYSDDHGPIKNVQVLVTETG
chr12:88900001210569 atggagaccgataaaaaaacatttaggcgcaaagcgcgagag
                         taaccgatttcccctctgtctgaagaaagctaatatttcacg
                         tttcaaagtgtaaatacagaatccttttaaaataagtgaaaa

AF063249_1          1800                           QQDGNVTKWYDAYFNKARPYFTN
                                                   Q DGNVTKWYDAYFNKARPYFTN
                                  A:A[gct]         QHDGNVTKWYDAYFNKARPYFTN
chr12:88900001210695 GGTATCAT  Intron 30  CAGCTccggagaattggttaagacttaa
                          <1-----[210696:21130-1>  aaagatcagaacataacgcatca
                                                   gttataaggttatttaagattat
```

Fig. 14 (continued)

```
AF063249_1        1824 EGFPNPPCIEGKTKFSGNEEIYVIGADNACMIPGNEEKICNGPLKPKKQ
                       EGFPNPPC EGKTKFSGNEEIY+IGADNACMIPGNE+KICNGPLKPKKQ
                       EGFPNPPCTEGKTKFSGNEEIYIIGADNACMIPGNEDKICNGPLKPKKQ
chr12:88900001211378 ggtcacctaggaaatagaggataagggagtaacgaggaatagccacaac
                       agtcaccgcagacatggaaatattgcaacgttcgaaaatgagctacaaa
                       acttctataaagagttctaacccattttacgttctacatctaagaaaga

AF063249_1        1873 Y                            FKFRATNVMGQFTDSEYSDPIK
                       Y                            FKFRATN+MGQFTDS+YSDP+K
                       Y            L:L[tta]        FKFRATNIMGQFTDSDYSDPVK
chr12:88900001211525 tTTGTAAGTA  Intron 31  TAGAtatagaaaagctagtgttgcga
                       a  <2-----[211530:21321-2> tatgccattgatcacaacacta
                       c                            tatatattgaattcttttcttg

AF063249_1        1897 TL                           EGLSERTVEIILSVTLCILSI
                       TL                           EGLSERTVEIILSVTLCILSI
                       TL           G:G[ggg]        EGLSERTVEIILSVTLCILSI
chr12:88900001213281 atGGTAAGAC  Intron 32  CAGGGggctgaaggaactgattacta
                       ct <1-----[213288:22334-1>  agtcagctatttctctgttct
                       ta                           aataaacagcttcctgtctaa

AF063249_1        1921 ILLGTAIFAFV                          IRQKQKEGGTYS
                       ILLGTAIFAF                           IRQKQKEGGTYS
                       ILLGTAIFAFA           R:R[aga]       IRQKQKEGGTYS
chr12:88900001223410 accgagatgtgAGGTAAGAT  Intron 33  TAGAaccacagggatt
                       tttgccttctc  <2-----[223445:22609-2> tgaaaaaggcac
                       tctaatttata                          tagggaatcact

AF063249_1        1945 PRDAEIIDTKFKLDQLITVADLELKDERLT
                       P+DAEIIDTK+KLDQLITVADLELKDERLT
                       PQDAEIIDTKLKLDQLITVADLELKDERLT
chr12:88900001226135 ccgggaagaatacgccaagggcgcaggata
                       caacattacatataattctcatataaagtc
                       tgtaattctagggtgccagacgaggcgaag

AF063249_1        1975                         LLSYRKSI
                                               LLSYRKSI
                                R:R[cgg]       LLSYRKSI
chr12:88900001226225 CGGTGAGCA  Intron 34  CAGGtcataata
                          <2-----[226227:24004-2> ttgagact
                                               atttaacc

AF063249_1        1984                         PISKKSFLQHVEELCTNSNLKFQ
                                               PISKKSFLQHVEELCTN+NLKFQ
                                K:K[aag]       PISKKSFLQHVEELCTNNNLKFQ
chr12:88900001240075 AAGTAAGTT  Intron 35  GAGGcaaaattcccgggctaaaacatc
                          <2-----[240077:24074-2> ctgaacttaataatgcaaatata
                                               aacgaccgattagtcacccagta
```

Fig. 14 (continued)

```
AF063249_1        2008 EEFS                            ELPKFLQDLSSTDADLPWNRAK
                       EEFS                            ELPKFLQDLSSTDADLPWNRAK
                       EEFS                            ELPKFLQDLSSTDADLPWNRAK
chr12:88900001240811 ggttGTATGTT  Intron 36  CAGgtcatccgcttagggcctaaga
                     aatc<0-----[240823:24389-0>atcattaatcccacatcgagca
                     aatg                              aaaattgtttattttgtgtaaa

AF063249_1        2034 NRFPNIKP                                NNNRVKLIADVSLPG
                       NRFPNIKP                                NNNRVKLIAD.S+PG
                       NRFPNIKP            Y:Y[tat]            NNNRVKLIADASVPG
chr12:88900001243963 actcaaacTGTATGTG  Intron 37  TAGATaaaagacagggagcg
                     agtcatac <1-----[243988:24912-1>  aaagtattcacgtcg
                     cccacaaa                                ttcaaggatctttat

AF063249_1        2058 SDYINASYVS                            GYLCPNEFIATQGPLP
                       SDYINASY+S                            GYLCPNEFIATQGPLP
                       SDYINASYIS                            GYLCPNEFIATQGPLP
chr12:88900001249173 tgtaagatatGTAAGTT  Intron 38  TAGgtttcagtagacgccc
                     caatacgatc<0-----[249203:26014-0>gatgcaattccagctc
                     gttttccttt                              ttatatattttataaa

AF063249_1        2084 GTVGDFWRMVWETRTKTLVMLTQCFEKGR
                       GTVGDFW+MVWETR KTLVMLTQCFEKGR
                       GTVGDFWKMVWETRAKTLVMLTQCFEKGR
chr12:88900001260196 gagggttaagtgaagaatgacacttgagcGTAAGTT  Intron 39
                     gctgatgattgacgcacttttcagtaagg<0-----[260283:260479
                     aatattgagggacaaaaaagaagttaaag

AF063249_1        2113    IRCHQYWPEDNKPVTVFGDIVITKLMEDIQIDWTIRDLKIER
                          IRCHQYWPEDNKPVTVFGDIVITKLMED+QIDWTIRDLKIER
                          IRCHQYWPEDNKPVTVFGDIVITKLMEDVQIDWTIRDLKIER
chr12:88900001260477 CAGaatccttcggaacgagtggagaaacagggcagtaaagcaaga
                     -0>tggaaagcaaaactcttgatttcattaatatagctgatatag
                        cactgtgagccgattctatagtagaggttaatgtcgtgatag

AF063249_1        2155                           HGDCMTVRQCNFTGWPEHGVPENTTP
                                                 HGDCMTVRQCNFT WPEHGVPEN+ P
                                                 HGDCMTVRQCNFTAWPEHGVPENSAP
chr12:88900001260606 GTAAAAA  Intron 40  TAGcggtaagcctatagtcgcggcgaagc
                     <0-----[260606:26148-0>agagtctgagatccgcaagtcaagcc
                                                 tgtcgttagtcttcgagtgttgccct

AF063249_1        2181 LIHFVKLVRTSRAHDTTPMVVHC
                       LIHFVKLVR SRAHDTTPM+VHC
                       LIHFVKLVRASRAHDTTPMIVHC             S:S[agt]
chr12:88900001261563 cactgatgcgaagcgaacaagctAGGTGAGAA  Intron 41  TAGT
                     ttattattgcggcaaccctttag  <2-----[261634:26430-2>
                     atctgggtaacgatccatgttcc
```

Fig. 14 (continued)

```
AF063249_1        2205 AGVGRTGVFIALDHLTQHINNHDFVDIYGLVAELRSERMCMVQNL
                       AGVGRTGVFIALDHLTQHIN+HDFVDIYGLVAELRSERMCMVQNL
                       AGVGRTGVFIALDHLTQHINDHDFVDIYGLVAELRSERMCMVQNL
chr12:88900001264311 ggggaaggtagcgctaccaagcgtggatgcgggcaagaatagcac
                       cgtggcgtttctaatcaataaaattatagttcatggagtgttaat
                       tataatattttgctaaatatctttgtataaatagataagcgggtg

AF063249_1        2250                            AQYIFLHQCILDLLSNKGGHQPVCFV
                                                  AQYIFLHQCILDLLSNKG +QP+CFV
                                                  AQYIFLHQCILDLLSNKGSNQPICFV
chr12:88900001264446 GTAAGAT  Intron 42  TAGgctattcctacgcttaagaaccattg
                       <0-----[264446:26974-0>caatttaagttattcaaggaactgtt
                                                  agtctacgctgtcaatgattgccttt

AF063249_1        2276 NYSTLQKMDSLDAME                          DVELEWEE
                       NYS LQKMDSLDAME                          DVELEWEE
                       NYSALQKMDSLDAME        G:G[ggt]          DVELEWEE
chr12:8890000------- attgccaagttggagGGTAAACA  Intron 43  TAGGTgggcgtgg
                       aacctaatactacta <1-----[314473:31473-1>  atatagaa
                       ctaatgggctgccga                          ttgtagaa

AF063249          2300 TTM
                       TTM
                       TTM
chr12:8890000------- aaa
                       cct
                       ctg
```

Fig. 15

```
492_BAC_p        1 ENSESFLWSTTSPSPTLGRVTPTVRTTQSSSTAARSKISSVWKEPISFV
                   ENSESFLWST SPSPTLGRVTP  RTT SSST   +++ISSVWKEPISFV
                   ENSESFLWSTASPSPTLGRVTPPSRTTHSSSTLTQNEISSVWKEPISFV
BAC          29570 gaagttttaagactcacgagacctcaacttaatacagaatgtagcaatg
                   aagacttggccgcccctggtccccgccaccgctcaaatgctgaactgtt
                   gttattagtaccttacttatatagtcataacggagtgcctggagtctta

492_BAC_p       50 VTHLRPYTTYLFEVSAVTTEAGYIDSTIVRTPES
                   VTHLRPYTTYLFEVSA TTEAGYIDSTIVRTPES
                   VTHLRPYTTYLFEVSAATTEAGYIDSTIVRTPES
BAC          29717 gactactaatctggtggaagggtagaaagaacgt
                   tcatgcaccattatcccccacgatagcttgccac
                   gacgattaatttatattataatttttgtcaaaaa

492_BAC_p       84 VPEGPPQNCIMGNVTGKAFSISWDPPTIVTGKFSYRVELYGPS
                   VPEGPPQNC+ GN+TGK+FSI WDPPTIVTGKFSYRVELYGPS
                   VPEGPPQNCVTGNITGKSFSILWDPPTIVTGKFSYRVELYGPS
BAC         -16606 gcggcccatgagaaagatttattgccaagagatataggttgct
                   tcagccaagtcgatcgactcttgacccttcgatgagtatagcc
                   gtaaaaaccaacccaagctatagcaataaagatttataataaa

492_BAC_p      127                           RILDNSTKDLRFAFTHLTPFTMY
                                             RILDNSTKDL+FAFT+LTPFTMY
                           G:G[ggt]          RILDNSTKDLKFAFTNLTPFTMY
BAC         -16477 GGTAAGCC  Intron 1   CAGGTcatgaaaagcatgtaacactaat
                    <1-----[16476:13211]-1>  gttaagcaatatctcatcctcta
                                             ctgtccaaccgtactcaaatagt

492_BAC_p      151 DVYVAAETSAGVGPKSNLSVFTPPD
                   DVY+AAETSAG GPKSN+SVFTPPD
                   DVYIAAETSAGTGPKSNISVFTPPD        V:V[gtt]
BAC         -13139 ggtagggaaggagcataatgtaccgGGTAAGAA  Intron 2
                   atatccacgcgcgcacatcttccca <1-----[13063:  883]
                   tcttgtactagtgcgattaactaat

492_BAC_p      176       PGAVFDLQIAEVEATEIRITWRKPRQPNGIISQYRVKVSVLETG
                         PGAVFDLQ+AEVE+T++RITW+KPRQPNGII+QYRVKV V ETG
                         PGAVFDLQLAEVESTQVRITWKKPRQPNGIINQYRVKVLVPETG
BAC           -885 TAGTTcgggtgtccggggtacgaaataaccccagaaactcgagcgcgag
                   -1>   cgcttatatcataccatgtcgaacgacagttaaagtatttcacg
                         agagttaatagaacgaaattggaaaaatattcacagagatagaa

492_BAC_p      221 VVLENTLLTGQDE
                   ++LENTLLTG +E
                   IILENTLLTGNNE
BAC           -748 aatgaatcagaag
                   tttaacttcgaaa
                   atgattgctattg
```

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN RECEPTOR TYROSINE PHOSPHATASE

<130> Lio395

<150> USSN 60/329,329

<151> 2001-10-16

<160> 8

<170> PatentIn version 3.1

<210> 1

<211> 6897

<212> DNA

<213> Homo sapiens

<400> 1
atggattttc ttatcatttt tcttttactt tttattggga cttcagagac acaggttgat 60

gtttccaatg tcgttcctgg tactaggtac gatataacca tctcttcaat ttctacaaca 120

tacacctcac ctgttactag aatagtgaca acaaatgtaa caaaaccagg gcctccagtc 180

ttcctagccg gggaaagagt cggatctgct gggattcttc tgtcttggaa tacaccacct 240

aatccaaatg gaaggattat atcttacatt gtcaaatata aggaagtttg tccgtggatg 300

caaacagtat atacacaagt cagatcaaag ccagacagtc tggaagttct tcttactaat 360

cttaatcctg gaacaacata tgaaattaag gttgctgctg aaaacagtgc tggcattgga 420

gtgtttagtg atccatttct cttccaaact gcagaaagtg ctccaggaaa agtggtgaat 480

ctcacagttg aggcctacaa cgcttcagca gttaagctga tttggtattt acctcggcaa 540

ccaaatggca aaattaccag cttcaagatt agtgtcaagc atgccagaag tgggatagta 600
gtgaaagatg tctcaatcag agtagaggac attttgactg ggaaattgcc agaaagcaat 660
gagaatagtg aatctttttt atggagtaca gccagccctt ctccaaccct tggtagagtt 720
acacctccat cgcgtaccac acattcatca agcacgttga cacagaatga gatcagctct 780
gtgtggaaag agcctatcag ttttgtagtg acacacttga gaccttatac aacatatctt 840
tttgaagttt cagctgctac aactgaagca ggttatattg atagtacgat tgtcagaaca 900
ccagaatcag tgcctgaagg accaccacaa aactgcgtaa caggcaacat cacaggaaag 960
tccttttcaa ttttatggga cccaccaact atagtaacag ggaaatttag ttatagagtt 1020
gaattatatg gaccatcagg tcgcattttg gataacagca caaaagacct caagtttgca 1080
ttcactaacc taacaccatt tacaatgtat gatgtctata ttgcggctga aaccagtgca 1140
gggactgggc ccaagtcaaa tatttcagta ttcactccac cagatgttcc aggggcagtg 1200
tttgatttac aacttgcaga ggtagaatcc acgcaagtaa gaattacttg gaagaaacca 1260
cgacaaccaa atggaattat taaccaatac cgagtgaaag tgctagttcc agagacagga 1320
ataattttgg aaaatacttt gctcactgga aataatgagt atataaatga ccccatggct 1380
ccagaaattg tgaacatagt agagccaatg gtaggattat atgagggttc agcagagatg 1440
tcgtctgacc ttcactcact tgctacattt atatataaca gccatccaga taaaaacttt 1500
cctgcaagga atagagctga agaccagact tcaccagttg taactacaag gaatcagtat 1560
attactgaca ttgcagctga acagctgtct tatgttatca ggagacttgt acctttcact 1620
gagcacatga ttagtgtatc tgctttcacc atcatgggag aaggaccacc aacagttctc 1680
agtgttagga cacgtcagca agtgccaagc tccattaaaa ttataaacta taaaaatatt 1740
agttcttcat ctattttgtt atattgggat cctccagaat atcccaatgg aaaaataact 1800
cactatacga tttatgcaat ggaattggat acaaacagag cattccagat aactaccata 1860
gataacagct ttctcataac agggttaaag aaatacacaa aatacaaaat gagagtggca 1920
gcctcaaccc acgttggaga aagttctttg tctgaagaaa atgacatctt tgtgagaact 1980
tcagaagatg aaccggaatc atcacctcaa gatgtcgaag taattgatgt taccgcagat 2040
gaaataaggt tgaagtggtc accacccgaa aagcccaatg ggatcattat tgcttatgaa 2100
gtgctatata aaaatataga tactttatat atgaagaaca catcaacaac agacataata 2160
ttaaggaact taagacctca caccctctat aacatttctg taaggtctta caccagattt 2220
ggtcatggca atcaggtatc ttctttactc tctgtaagga cttcggagac tgtgcctgat 2280
agtgcaccag aaaatatcac ttacaaaaat atttcttctg gagagattga gctatcattc 2340

cttcccccaa gtagtcccaa tggaatcata caaaaatata caatttatct caagagaagt 2400
aatggaaatg aggaaagaac tataaataca acctctttaa cccaaaacat taaagtactg 2460
aagaaatata cccaatatat cattgaggtg tctgctagta cactcaaagg tgaaggagtt 2520
cggagtgctc ccataagtat actgacggag gaagatgctc ctgattctcc ccctcaagac 2580
ttctctgtaa aacagttgtc tggtgtcacg gtgaagttgt catggcaacc acccctggag 2640
ccaaatggaa ttatccttta ttacacagtt tatgtctgga atagatcatc attaaaaact 2700
attaatgtca ctgaaacatc attggagtta tcagatttgg attataatgt tgaatacagt 2760
gcttatgtaa cagctagcac cagatttggt gatgggaaaa caagaagcaa tatcattagc 2820
tttcaaacac cagagggagc accaagcgat cctcccaaag atgtttatta tgcaaacctc 2880
agttcttcat caataattct tttctggaca cctccttcaa aacctaatgg gattatacaa 2940
tattactctg tttattacag aaatacttca ggtactttta tgcagaattt tacactccat 3000
gaagtaacca atgactttga caatatgact gtatccacaa ttatagataa actgacaata 3060
ttcagctact atacattttg gttaacagca agtacttcag ttggaaatgg gaataaaagc 3120
agtgacatca ttgaagtata cacagatcaa gacatacctg aagggtttgt tggaaacctg 3180
acttacgaat ccatttcgtc aactgcaata aatgtaagct gggtcccacc ggctcaacca 3240
aacggtctag tcttctacta tgtttcactg atcttacagc agactcctcg ccatgtgaga 3300
ccacctcttg ttacatatga gagaagcata tattttgata atctggaaaa atacactgat 3360
tatatattaa aaattactcc atcaacagaa aagggattct ctgataccta tactgcccag 3420
ctatacatca agactgaaga agatgtccca gaaacttcac caataatcaa cacttttaaa 3480
aacctttcct ctacctcagt tctcttatca tgggatcccc cagtaaagcc aaatggtgca 3540
ataataagtt atgatttaac tttacaagga ccaaatgaaa attattcttt cattacttct 3600
gataattaca taatattgga agagctttca ccatttacat tatatagctt ttttgctgcc 3660
gcaagaacta gaaaaggact tggtccttcc agtattcttt tcttttacac agatgagtca 3720
gtgccgttag cacctccaca aaatttgact ttaatcaact gtacttcaga ctttgtatgg 3780
ctgaaatgga gcccaagtcc tcttccaggt ggtattgtta aagtatatag ttttaaaatt 3840
catgaacatg aaactgacac tatatattat aagaatatat caggatttaa aactgaagcc 3900
aaacttgttg gactggaacc agtcagcacc tactctatcc gtgtatctgc gttcaccaaa 3960
gttggaaatg gcaatcaatt tagtaatgta gtaaaattca caacccaaga atcagttcca 4020
gatgtcgtgc agaatatgca gtgcatggca actagctggc agtcagtttt agtgaaatgg 4080

gatccaccca aaaaggcaaa tggaataata acgcagtata tggtaacagt tgaaaggaat 4140
tctacaaaag tttctcccca agatcacatg tacactttca taaagcttct tgccaatacc 4200
tcatatgtct ttaaagtaag agcttcaacc tcagctggtg aaggtgatga aagcacatgc 4260
catgtcagca cactacctga aacagttccc agtgttccca caaatattgc tttttctgat 4320
gttcagtcaa ctagtgcaac attgacatgg ataagacctg acactatcct tggctacttt 4380
caaaattaca aaattaccac tcaacttcgt gctcaaaaat gcaaagaatg ggaatccgaa 4440
gaatgtgttg aatatcaaaa aattcaatac ctctatgaag ctcacttaac tgaagagaca 4500
gtatatggat taaagaaatt tagatggtat agattccaag tggctgccag caccaatgct 4560
ggctatggca atgcttcaaa ctggatttct acaaaaactc tgcctggccc tccagatggt 4620
cctcctgaaa atgttcatgt agtagcaaca tcacctttta gcatcagcat aagctggagt 4680
gaacctgctg tcattactgg accaacatgt tatctgattg atgtcaaatc ggtagataat 4740
gatgaattta atatatcctt catcaagtca aatgaagaaa ataaaaccat agaaattaaa 4800
gatttagaaa tattcacaag gtattctgta gtgatcactg catttactgg gaacattagt 4860
gctgcatatg tagaagggaa gtcaagtgct gaaatgattg ttactacttt agaatcagcc 4920
ccaaaggacc cacctaacaa catgacattt cagaagatac cagatgaagt tacaaaattt 4980
caattaacgt tccttcctcc ttctcaacct aatggaaata tccaagtata tcaagctctg 5040
gtttaccgag aagatgatcc tactgctgtc cagattcaca acctcagtat tatacagaaa 5100
accaacacat tcgtcattgc aatgctagaa ggactaaaag gtggacatac atacaatatc 5160
agtgtttacg cagtcaatag tgctggtgca ggtccaaagg ttccgatgag aataaccatg 5220
gatatcaaag ctccagcacg accaaaaacc aaaccaaccc ctatttatga tgccacagga 5280
aaactgcttg tgacttcaac aacaattaca atcagaatgc caatatgtta ctacagtgat 5340
gatcatggac caataaaaaa tgtacaagtg cttgtgacag aaacaggagc tcagcatgat 5400
ggaaatgtaa caaagtggta tgatgcatat tttaataaag caaggccata ttttacaaat 5460
gaaggctttc ctaaccctcc atgtacagaa ggaaagacaa agtttagtgg caatgaagaa 5520
atctacatca taggtgctga taatgcatgc atgattcctg gcaatgaaga caaaatttgc 5580
aatggaccac tgaaaccaaa aaagcaatac ttatttaaat ttagagctac aaatattatg 5640
ggacaattta ctgactctga ttattctgac cctgttaaga ctttagggga aggactttca 5700
gaaagaaccg tagagatcat tctttccgtc actttgtgta tcctttcaat aattctcctt 5760
ggaacagcta tttttgcatt tgcaagaatt cgacagaagc agaaagaagg tggcacatac 5820
tctcctcagg atgcagaaat tattgacact aaattgaagc tggatcagct catcacagtg 5880

gcagacctgg aactgaagga cgagagatta acgcggttac ttagttatag aaaatccatc 5940
aagccaataa gcaagaaatc cttcctgcaa catgttgaag agctttgcac aaacaacaac 6000
ctaaagtttc aagaagaatt ttcggaatta ccaaaatttc ttcaggatct ttcttcaact 6060
gatgctgatc tgccttggaa tagagcaaaa aaccgcttcc caaacataaa accatataat 6120
aataacagag taaagctgat agctgacgct agtgttccag gttcggatta tattaatgcc 6180
agctatattt ctggttattt atgtccaaat gaatttattg ctactcaagg tccactacca 6240
ggaacagttg gagatttttg gaaaatggtg tgggaaacca gagcaaaaac attagtaatg 6300
ctaacacagt gttttgaaaa aggacggatc agatgccatc agtattggcc agaggacaac 6360
aagccagtta ctgtctttgg agatatagtg attacaaagc taatggagga tgttcaaata 6420
gattggacta tcagggatct gaaaattgaa aggcatgggg attgcatgac tgttcgacag 6480
tgtaacttta ctgcctggcc agagcatggg gttcctgaga acagcgcccc tctaattcac 6540
tttgtgaagt tggttcgagc aagcagggca catgacacca cacctatgat tgttcactgc 6600
agtgctggag ttggaagaac tggagttttt attgctctgg accatttaac acaacatata 6660
aatgaccatg attttgtgga tatatatgga ctagtagctg aactgagaag tgaaagaatg 6720
tgcatggtgc agaatctggc acagtatatc tttttacacc agtgcattct ggatctctta 6780
tcaaataagg gaagtaatca gcccatctgt tttgttaact attcagcact tcagaagatg 6840
gactctttgg acgccatgga aggtgatgtt gagcttgaat gggaagaaac cactatg 6897

<210> 2

<211> 2299

<212> PRT

<213> Homo sapiens

<400> 2

Met Asp Phe Leu Ile Ile Phe Leu Leu Leu Phe Ile Gly Thr Ser Glu
1 5 10 15

Thr Gln Val Asp Val Ser Asn Val Val Pro Gly Thr Arg Tyr Asp Ile
20 25 30

Thr Ile Ser Ser Ile Ser Thr Thr Tyr Thr Ser Pro Val Thr Arg Ile
35 40 45

Val Thr Thr Asn Val Thr Lys Pro Gly Pro Pro Val Phe Leu Ala Gly
50 55 60

Glu Arg Val Gly Ser Ala Gly Ile Leu Leu Ser Trp Asn Thr Pro Pro
65 70 75 80

Asn Pro Asn Gly Arg Ile Ile Ser Tyr Ile Val Lys Tyr Lys Glu Val
85 90 95

Cys Pro Trp Met Gln Thr Val Tyr Thr Gln Val Arg Ser Lys Pro Asp
100 105 110

Ser Leu Glu Val Leu Leu Thr Asn Leu Asn Pro Gly Thr Thr Tyr Glu
115 120 125

Ile Lys Val Ala Ala Glu Asn Ser Ala Gly Ile Gly Val Phe Ser Asp
130 135 140

Pro Phe Leu Phe Gln Thr Ala Glu Ser Ala Pro Gly Lys Val Val Asn
145 150 155 160

Leu Thr Val Glu Ala Tyr Asn Ala Ser Ala Val Lys Leu Ile Trp Tyr
165 170 175

Leu Pro Arg Gln Pro Asn Gly Lys Ile Thr Ser Phe Lys Ile Ser Val
180 185 190

Lys His Ala Arg Ser Gly Ile Val Val Lys Asp Val Ser Ile Arg Val
195 200 205

Glu Asp Ile Leu Thr Gly Lys Leu Pro Glu Ser Asn Glu Asn Ser Glu
210 215 220

Ser Phe Leu Trp Ser Thr Ala Ser Pro Ser Pro Thr Leu Gly Arg Val
225 230 235 240

Thr Pro Pro Ser Arg Thr Thr His Ser Ser Ser Thr Leu Thr Gln Asn
245 250 255

Glu Ile Ser Ser Val Trp Lys Glu Pro Ile Ser Phe Val Val Thr His
260 265 270

Leu Arg Pro Tyr Thr Thr Tyr Leu Phe Glu Val Ser Ala Ala Thr Thr
275 280 285

Glu Ala Gly Tyr Ile Asp Ser Thr Ile Val Arg Thr Pro Glu Ser Val
290 295 300

Pro Glu Gly Pro Pro Gln Asn Cys Val Thr Gly Asn Ile Thr Gly Lys
305 310 315 320

Ser Phe Ser Ile Leu Trp Asp Pro Pro Thr Ile Val Thr Gly Lys Phe
325 330 335

Ser Tyr Arg Val Glu Leu Tyr Gly Pro Ser Gly Arg Ile Leu Asp Asn
340 345 350

Ser Thr Lys Asp Leu Lys Phe Ala Phe Thr Asn Leu Thr Pro Phe Thr
355 360 365

Met Tyr Asp Val Tyr Ile Ala Ala Glu Thr Ser Ala Gly Thr Gly Pro
370 375 380

Lys Ser Asn Ile Ser Val Phe Thr Pro Pro Asp Val Pro Gly Ala Val
385 390 395 400

Phe Asp Leu Gln Leu Ala Glu Val Glu Ser Thr Gln Val Arg Ile Thr
405 410 415

Trp Lys Lys Pro Arg Gln Pro Asn Gly Ile Ile Asn Gln Tyr Arg Val
420 425 430

Lys Val Leu Val Pro Glu Thr Gly Ile Ile Leu Glu Asn Thr Leu Leu
435 440 445

Thr Gly Asn Asn Glu Tyr Ile Asn Asp Pro Met Ala Pro Glu Ile Val
450 455 460

Asn Ile Val Glu Pro Met Val Gly Leu Tyr Glu Gly Ser Ala Glu Met
465 470 475 480

Ser Ser Asp Leu His Ser Leu Ala Thr Phe Ile Tyr Asn Ser His Pro
485 490 495

Asp Lys Asn Phe Pro Ala Arg Asn Arg Ala Glu Asp Gln Thr Ser Pro
500 505 510

Val Val Thr Thr Arg Asn Gln Tyr Ile Thr Asp Ile Ala Ala Glu Gln
515 520 525

Leu Ser Tyr Val Ile Arg Arg Leu Val Pro Phe Thr Glu His Met Ile
530 535 540

Ser Val Ser Ala Phe Thr Ile Met Gly Glu Gly Pro Pro Thr Val Leu
545 550 555 560

Ser Val Arg Thr Arg Gln Gln Val Pro Ser Ser Ile Lys Ile Ile Asn
565 570 575

Tyr Lys Asn Ile Ser Ser Ser Ser Ile Leu Leu Tyr Trp Asp Pro Pro
580 585 590

Glu Tyr Pro Asn Gly Lys Ile Thr His Tyr Thr Ile Tyr Ala Met Glu
595 600 605

Leu Asp Thr Asn Arg Ala Phe Gln Ile Thr Thr Ile Asp Asn Ser Phe
610 615 620

Leu Ile Thr Gly Leu Lys Lys Tyr Thr Lys Tyr Lys Met Arg Val Ala
625 630 635 640

Ala Ser Thr His Val Gly Glu Ser Ser Leu Ser Glu Glu Asn Asp Ile
645 650 655

Phe Val Arg Thr Ser Glu Asp Glu Pro Glu Ser Ser Pro Gln Asp Val
660 665 670

Glu Val Ile Asp Val Thr Ala Asp Glu Ile Arg Leu Lys Trp Ser Pro
675 680 685

Pro Glu Lys Pro Asn Gly Ile Ile Ile Ala Tyr Glu Val Leu Tyr Lys
690 695 700

Asn Ile Asp Thr Leu Tyr Met Lys Asn Thr Ser Thr Thr Asp Ile Ile
705 710 715 720

Leu Arg Asn Leu Arg Pro His Thr Leu Tyr Asn Ile Ser Val Arg Ser
725 730 735

Tyr Thr Arg Phe Gly His Gly Asn Gln Val Ser Ser Leu Leu Ser Val
740 745 750

Arg Thr Ser Glu Thr Val Pro Asp Ser Ala Pro Glu Asn Ile Thr Tyr
755 760 765

Lys Asn Ile Ser Ser Gly Glu Ile Glu Leu Ser Phe Leu Pro Pro Ser
770 775 780

Ser Pro Asn Gly Ile Ile Gln Lys Tyr Thr Ile Tyr Leu Lys Arg Ser
785 790 795 800

Asn Gly Asn Glu Glu Arg Thr Ile Asn Thr Thr Ser Leu Thr Gln Asn
805 810 815

Ile Lys Val Leu Lys Lys Tyr Thr Gln Tyr Ile Ile Glu Val Ser Ala
820 825 830

Ser Thr Leu Lys Gly Glu Gly Val Arg Ser Ala Pro Ile Ser Ile Leu
835 840 845

Thr Glu Glu Asp Ala Pro Asp Ser Pro Pro Gln Asp Phe Ser Val Lys
850 855 860

Gln Leu Ser Gly Val Thr Val Lys Leu Ser Trp Gln Pro Pro Leu Glu
865 870 875 880

Pro Asn Gly Ile Ile Leu Tyr Tyr Thr Val Tyr Val Trp Asn Arg Ser
885 890 895

Ser Leu Lys Thr Ile Asn Val Thr Glu Thr Ser Leu Glu Leu Ser Asp
900 905 910

Leu Asp Tyr Asn Val Glu Tyr Ser Ala Tyr Val Thr Ala Ser Thr Arg
915 920 925

Phe Gly Asp Gly Lys Thr Arg Ser Asn Ile Ile Ser Phe Gln Thr Pro
930 935 940

Glu Gly Ala Pro Ser Asp Pro Pro Lys Asp Val Tyr Tyr Ala Asn Leu
945 950 955 960

Ser Ser Ser Ser Ile Ile Leu Phe Trp Thr Pro Pro Ser Lys Pro Asn
965 970 975

Gly Ile Ile Gln Tyr Tyr Ser Val Tyr Tyr Arg Asn Thr Ser Gly Thr
980 985 990

Phe Met Gln Asn Phe Thr Leu His Glu Val Thr Asn Asp Phe Asp Asn
995 1000 1005

Met Thr Val Ser Thr Ile Ile Asp Lys Leu Thr Ile Phe Ser Tyr
1010 1015 1020

Tyr Thr Phe Trp Leu Thr Ala Ser Thr Ser Val Gly Asn Gly Asn
1025 1030 1035

Lys Ser Ser Asp Ile Ile Glu Val Tyr Thr Asp Gln Asp Ile Pro
1040 1045 1050

Glu Gly Phe Val Gly Asn Leu Thr Tyr Glu Ser Ile Ser Ser Thr
1055 1060 1065

Ala Ile Asn Val Ser Trp Val Pro Pro Ala Gln Pro Asn Gly Leu
1070 1075 1080

Val Phe Tyr Tyr Val Ser Leu Ile Leu Gln Gln Thr Pro Arg His
1085 1090 1095

Val Arg Pro Pro Leu Val Thr Tyr Glu Arg Ser Ile Tyr Phe Asp
1100 1105 1110

Asn Leu Glu Lys Tyr Thr Asp Tyr Ile Leu Lys Ile Thr Pro Ser
1115 1120 1125

Thr Glu Lys Gly Phe Ser Asp Thr Tyr Thr Ala Gln Leu Tyr Ile
1130 1135 1140

Lys Thr Glu Glu Asp Val Pro Glu Thr Ser Pro Ile Ile Asn Thr
1145 1150 1155

Phe Lys Asn Leu Ser Ser Thr Ser Val Leu Leu Ser Trp Asp Pro
1160 1165 1170

Pro Val Lys Pro Asn Gly Ala Ile Ile Ser Tyr Asp Leu Thr Leu
1175 1180 1185

Gln Gly Pro Asn Glu Asn Tyr Ser Phe Ile Thr Ser Asp Asn Tyr
1190 1195 1200

Ile Ile Leu Glu Glu Leu Ser Pro Phe Thr Leu Tyr Ser Phe Phe
1205 1210 1215

Ala Ala Ala Arg Thr Arg Lys Gly Leu Gly Pro Ser Ser Ile Leu
1220 1225 1230

Phe Phe Tyr Thr Asp Glu Ser Val Pro Leu Ala Pro Pro Gln Asn
1235 1240 1245

Leu Thr Leu Ile Asn Cys Thr Ser Asp Phe Val Trp Leu Lys Trp
1250 1255 1260

Ser Pro Ser Pro Leu Pro Gly Gly Ile Val Lys Val Tyr Ser Phe
1265 1270 1275

Lys Ile His Glu His Glu Thr Asp Thr Ile Tyr Tyr Lys Asn Ile
1280 1285 1290

Ser Gly Phe Lys Thr Glu Ala Lys Leu Val Gly Leu Glu Pro Val
1295 1300 1305

Ser Thr Tyr Ser Ile Arg Val Ser Ala Phe Thr Lys Val Gly Asn
1310 1315 1320

Gly Asn Gln Phe Ser Asn Val Val Lys Phe Thr Thr Gln Glu Ser
1325 1330 1335

Val Pro Asp Val Val Gln Asn Met Gln Cys Met Ala Thr Ser Trp
1340 1345 1350

Gln Ser Val Leu Val Lys Trp Asp Pro Pro Lys Lys Ala Asn Gly
1355 1360 1365

Ile Ile Thr Gln Tyr Met Val Thr Val Glu Arg Asn Ser Thr Lys
1370 1375 1380

Val Ser Pro Gln Asp His Met Tyr Thr Phe Ile Lys Leu Leu Ala
1385 1390 1395

Asn Thr Ser Tyr Val Phe Lys Val Arg Ala Ser Thr Ser Ala Gly
1400 1405 1410

Glu Gly Asp Glu Ser Thr Cys His Val Ser Thr Leu Pro Glu Thr
1415 1420 1425

Val Pro Ser Val Pro Thr Asn Ile Ala Phe Ser Asp Val Gln Ser
1430 1435 1440

Thr Ser Ala Thr Leu Thr Trp Ile Arg Pro Asp Thr Ile Leu Gly
1445 1450 1455

Tyr Phe Gln Asn Tyr Lys Ile Thr Thr Gln Leu Arg Ala Gln Lys
1460 1465 1470

Cys Lys Glu Trp Glu Ser Glu Glu Cys Val Glu Tyr Gln Lys Ile
1475 1480 1485

Gln Tyr Leu Tyr Glu Ala His Leu Thr Glu Glu Thr Val Tyr Gly
1490 1495 1500

Leu Lys Lys Phe Arg Trp Tyr Arg Phe Gln Val Ala Ala Ser Thr
1505 1510 1515

Asn Ala Gly Tyr Gly Asn Ala Ser Asn Trp Ile Ser Thr Lys Thr
1520 1525 1530

Leu Pro Gly Pro Pro Asp Gly Pro Pro Glu Asn Val His Val Val
1535 1540 1545

Ala Thr Ser Pro Phe Ser Ile Ser Ile Ser Trp Ser Glu Pro Ala
1550 1555 1560

Val Ile Thr Gly Pro Thr Cys Tyr Leu Ile Asp Val Lys Ser Val
1565 1570 1575

Asp Asn Asp Glu Phe Asn Ile Ser Phe Ile Lys Ser Asn Glu Glu
1580 1585 1590

Asn Lys Thr Ile Glu Ile Lys Asp Leu Glu Ile Phe Thr Arg Tyr
1595 1600 1605

er Val Val Ile Thr Ala Phe Thr Gly Asn Ile Ser Ala Ala Tyr
1610 1615 1620

ıl Glu Gly Lys Ser Ser Ala Glu Met Ile Val Thr Thr Leu Glu
1625 1630 1635

Ser Ala Pro Lys Asp Pro Pro Asn Asn Met Thr Phe Gln Lys Ile
1640 1645 1650

Pro Asp Glu Val Thr Lys Phe Gln Leu Thr Phe Leu Pro Pro Ser
1655 1660 1665

Gln Pro Asn Gly Asn Ile Gln Val Tyr Gln Ala Leu Val Tyr Arg
1670 1675 1680

Glu Asp Asp Pro Thr Ala Val Gln Ile His Asn Leu Ser Ile Ile
1685 1690 1695

Gln Lys Thr Asn Thr Phe Val Ile Ala Met Leu Glu Gly Leu Lys
1700 1705 1710

Gly Gly His Thr Tyr Asn Ile Ser Val Tyr Ala Val Asn Ser Ala
1715 1720 1725

Gly Ala Gly Pro Lys Val Pro Met Arg Ile Thr Met Asp Ile Lys
1730 1735 1740

Ala Pro Ala Arg Pro Lys Thr Lys Pro Thr Pro Ile Tyr Asp Ala
1745 1750 1755

Thr Gly Lys Leu Leu Val Thr Ser Thr Thr Ile Thr Ile Arg Met
1760 1765 1770

Pro Ile Cys Tyr Tyr Ser Asp Asp His Gly Pro Ile Lys Asn Val
1775 1780 1785

Gln Val Leu Val Thr Glu Thr Gly Ala Gln His Asp Gly Asn Val
1790 1795 1800

Thr Lys Trp Tyr Asp Ala Tyr Phe Asn Lys Ala Arg Pro Tyr Phe
1805 1810 1815

Thr Asn Glu Gly Phe Pro Asn Pro Pro Cys Thr Glu Gly Lys Thr
1820 1825 1830

Lys Phe Ser Gly Asn Glu Glu Ile Tyr Ile Ile Gly Ala Asp Asn
1835 1840 1845

Ala Cys Met Ile Pro Gly Asn Glu Asp Lys Ile Cys Asn Gly Pro
1850 1855 1860

Leu Lys Pro Lys Lys Gln Tyr Leu Phe Lys Phe Arg Ala Thr Asn
1865 1870 1875

Ile Met Gly Gln Phe Thr Asp Ser Asp Tyr Ser Asp Pro Val Lys
1880 1885 1890

Thr Leu Gly Glu Gly Leu Ser Glu Arg Thr Val Glu Ile Ile Leu
1895 1900 1905

Ser Val Thr Leu Cys Ile Leu Ser Ile Ile Leu Leu Gly Thr Ala
1910 1915 1920

Ile Phe Ala Phe Ala Arg Ile Arg Gln Lys Gln Lys Glu Gly Gly
1925 1930 1935

Thr Tyr Ser Pro Gln Asp Ala Glu Ile Ile Asp Thr Lys Leu Lys
1940 1945 1950

Leu Asp Gln Leu Ile Thr Val Ala Asp Leu Glu Leu Lys Asp Glu
1955 1960 1965

Arg Leu Thr Arg Leu Leu Ser Tyr Arg Lys Ser Ile Lys Pro Ile
1970 1975 1980

Ser Lys Lys Ser Phe Leu Gln His Val Glu Glu Leu Cys Thr Asn
1985 1990 1995

Asn Asn Leu Lys Phe Gln Glu Glu Phe Ser Glu Leu Pro Lys Phe
2000 2005 2010

Leu Gln Asp Leu Ser Ser Thr Asp Ala Asp Leu Pro Trp Asn Arg
2015 2020 2025

Ala Lys Asn Arg Phe Pro Asn Ile Lys Pro Tyr Asn Asn Asn Arg
2030 2035 2040

Val Lys Leu Ile Ala Asp Ala Ser Val Pro Gly Ser Asp Tyr Ile
2045 2050 2055

Asn Ala Ser Tyr Ile Ser Gly Tyr Leu Cys Pro Asn Glu Phe Ile
2060 2065 2070

Ala Thr Gln Gly Pro Leu Pro Gly Thr Val Gly Asp Phe Trp Lys
2075 2080 2085

Met Val Trp Glu Thr Arg Ala Lys Thr Leu Val Met Leu Thr Gln
2090 2095 2100

Cys Phe Glu Lys Gly Arg Ile Arg Cys His Gln Tyr Trp Pro Glu
2105 2110 2115

Asp Asn Lys Pro Val Thr Val Phe Gly Asp Ile Val Ile Thr Lys
2120 2125 2130

Leu Met Glu Asp Val Gln Ile Asp Trp Thr Ile Arg Asp Leu Lys
2135 2140 2145

Ile Glu Arg His Gly Asp Cys Met Thr Val Arg Gln Cys Asn Phe
2150 2155 2160

Thr Ala Trp Pro Glu His Gly Val Pro Glu Asn Ser Ala Pro Leu
2165 2170 2175

Ile His Phe Val Lys Leu Val Arg Ala Ser Arg Ala His Asp Thr
2180 2185 2190

Thr Pro Met Ile Val His Cys Ser Ala Gly Val Gly Arg Thr Gly
2195 2200 2205

Val Phe Ile Ala Leu Asp His Leu Thr Gln His Ile Asn Asp His
2210 2215 2220

Asp Phe Val Asp Ile Tyr Gly Leu Val Ala Glu Leu Arg Ser Glu
2225 2230 2235

Arg Met Cys Met Val Gln Asn Leu Ala Gln Tyr Ile Phe Leu His
2240 2245 2250

Gln Cys Ile Leu Asp Leu Leu Ser Asn Lys Gly Ser Asn Gln Pro
2255 2260 2265

Ile Cys Phe Val Asn Tyr Ser Ala Leu Gln Lys Met Asp Ser Leu
2270 2275 2280

Asp Ala Met Glu Gly Asp Val Glu Leu Glu Trp Glu Glu Thr Thr
2285 2290 2295

Met

<210> 3

<211> 2302

<212> PRT

<213> Homo sapiens

<400> 3

```
Met Met Asp Phe His Phe Ser Phe Leu Phe Leu Leu Ile Gly Thr Ser
1               5                   10                  15

Glu Ser Gln Val Asp Val Ser Ser Ser Phe Asp Gly Thr Gly Tyr Asp
            20                  25                  30

Ile Thr Leu Ser Ser Val Ser Ala Thr Thr Tyr Ser Ser Pro Val Ser
        35                  40                  45

Arg Thr Leu Ala Thr Asn Val Thr Lys Pro Gly Pro Pro Val Phe Leu
    50                  55                  60

Ala Gly Glu Arg Val Gly Ser Ala Gly Ile Leu Leu Ser Trp Asn Thr
65                  70                  75                  80

Pro Pro Asn Pro Asn Gly Arg Ile Ile Ser Tyr Val Val Lys Tyr Lys
                85                  90                  95

Glu Val Cys Pro Trp Met Gln Thr Ala Tyr Thr Arg Ala Arg Ala Lys
            100                 105                 110

Pro Asp Ser Leu Glu Val Leu Leu Thr Asn Leu Asn Pro Gly Thr Thr
        115                 120                 125

Tyr Glu Ile Lys Val Ala Ala Glu Asn Asn Ala Gly Ile Gly Val Phe
    130                 135                 140

Ser Asp Pro Phe Leu Phe Gln Thr Ala Glu Ser Ala Pro Gly Lys Val
145                 150                 155                 160

Val Asn Leu Thr Val Glu Ala Leu Asn Tyr Ser Ala Val Asn Leu Ile
                165                 170                 175
```

```
Trp Tyr Leu Pro Arg Gln Pro Asn Gly Lys Ile Thr Ser Phe Lys Ile
            180                 185                 190

Ser Val Lys His Ala Arg Ser Gly Ile Val Val Lys Asp Val Ser Leu
        195                 200                 205

Arg Val Glu Asp Ile Leu Ser Gly Lys Leu Pro Glu Cys Asn Glu Asn
    210                 215                 220

Ser Glu Ser Phe Leu Trp Ser Thr Thr Ser Pro Ser Pro Thr Leu Gly
225                 230                 235                 240

Arg Val Thr Pro Thr Val Arg Thr Thr Gln Ser Ser Ser Thr Ala Ala
                245                 250                 255

Arg Ser Lys Ile Ser Ser Val Trp Lys Glu Pro Ile Ser Phe Val Val
            260                 265                 270

Thr His Leu Arg Pro Tyr Thr Thr Tyr Leu Phe Glu Val Ser Ala Val
        275                 280                 285

Thr Thr Glu Ala Gly Tyr Ile Asp Ser Thr Ile Val Arg Thr Pro Glu
    290                 295                 300

Ser Val Pro Glu Gly Pro Pro Gln Asn Cys Ile Met Gly Asn Val Thr
305                 310                 315                 320

Gly Lys Ala Phe Ser Ile Ser Trp Asp Pro Pro Thr Ile Val Thr Gly
                325                 330                 335

Lys Phe Ser Tyr Arg Val Glu Leu Tyr Gly Pro Ser Gly Arg Ile Leu
            340                 345                 350

Asp Asn Ser Thr Lys Asp Leu Arg Phe Ala Phe Thr His Leu Thr Pro
        355                 360                 365

Phe Thr Met Tyr Asp Val Tyr Val Ala Ala Glu Thr Ser Ala Gly Val
    370                 375                 380

Gly Pro Lys Ser Asn Leu Ser Val Phe Thr Pro Pro Asp Val Pro Gly
385                 390                 395                 400
```

Ala Val Phe Asp Leu Gln Ile Ala Glu Val Glu Ala Thr Glu Ile Arg
405 410 415

Ile Thr Trp Arg Lys Pro Arg Gln Pro Asn Gly Ile Ile Ser Gln Tyr
420 425 430

Arg Val Lys Val Ser Val Leu Glu Thr Gly Val Val Leu Glu Asn Thr
435 440 445

Leu Leu Thr Gly Gln Asp Glu Ser Ile Ser Asn Pro Met Ser Pro Glu
450 455 460

Ile Met Asn Leu Val Asp Pro Met Ile Gly Phe Tyr Glu Gly Ser Gly
465 470 475 480

Glu Met Ser Ser Asp Leu His Ser Pro Ala Ser Phe Ile Tyr Asn Ser
485 490 495

His Pro His Asn Asp Phe Pro Ala Ser Thr Arg Ala Glu Glu Gln Ser
500 505 510

Ser Pro Val Val Thr Thr Arg Asn Gln Tyr Met Thr Asp Ile Thr Ala
515 520 525

Glu Gln Leu Ser Tyr Val Val Arg Arg Leu Val Pro Phe Thr Glu His
530 535 540

Thr Ile Ser Val Ser Ala Phe Thr Ile Met Gly Glu Gly Pro Pro Thr
545 550 555 560

Val Leu Thr Val Arg Thr Arg Glu Gln Val Pro Ser Ser Ile Gln Ile
565 570 575

Ile Asn Tyr Lys Asn Ile Ser Ser Ser Ser Ile Leu Leu Tyr Trp Asp
580 585 590

Pro Pro Glu Tyr Pro Asn Gly Lys Ile Thr His Tyr Thr Ile Tyr Ala
595 600 605

Thr Glu Leu Asp Thr Asn Arg Ala Phe Gln Met Thr Thr Val Asp Asn
610 615 620

Ser Phe Leu Ile Thr Gly Leu Lys Lys Tyr Thr Arg Tyr Lys Met Arg
625 630 635 640

Val Ala Ala Ser Thr His Val Gly Glu Ser Ser Leu Ser Glu Glu Asn
645 650 655

Asp Ile Phe Val Arg Thr Pro Glu Asp Glu Pro Glu Ser Ser Pro Gln
660 665 670

Asp Val Gln Val Thr Gly Val Ser Pro Ser Glu Leu Arg Leu Lys Trp
675 680 685

Ser Pro Pro Glu Lys Pro Asn Gly Ile Ile Ile Ala Tyr Glu Val Leu
690 695 700

Tyr Gln Asn Ala Asp Thr Leu Phe Val Lys Asn Thr Ser Thr Thr Asp
705 710 715 720

Ile Ile Ile Ser Asp Leu Lys Pro Tyr Thr Leu Tyr Asn Ile Ser Ile
725 730 735

Arg Ser Tyr Thr Arg Leu Gly His Gly Asn Gln Ser Ser Ser Leu Leu
740 745 750

Ser Val Arg Thr Ser Glu Thr Val Pro Asp Ser Ala Pro Glu Asn Ile
755 760 765

Thr Tyr Lys Asn Ile Ser Ser Gly Glu Ile Glu Ile Ser Phe Leu Pro
770 775 780

Pro Arg Ser Pro Asn Gly Ile Ile Gln Lys Tyr Thr Ile Tyr Leu Lys
785 790 795 800

Arg Ser Asn Ser His Glu Ala Arg Thr Ile Asn Thr Thr Ser Leu Thr
805 810 815

Gln Thr Ile Gly Gly Leu Lys Lys Tyr Thr His Tyr Val Ile Glu Val
820 825 830

Ser Ala Ser Thr Leu Lys Gly Glu Gly Ile Arg Ser Arg Pro Ile Ser
835 840 845

Ile Leu Thr Glu Glu Asp Ala Pro Asp Ser Pro Pro Gln Asn Phe Ser
850 855 860

Val Lys Gln Leu Ser Gly Val Thr Val Met Leu Ser Trp Gln Pro Pro
865 870 875 880

Leu Glu Pro Asn Gly Ile Ile Leu Tyr Tyr Thr Val Tyr Val Trp Asp
885 890 895

Lys Ser Ser Leu Arg Ala Ile Asn Ala Thr Glu Ala Ser Leu Val Leu
900 905 910

Ser Asp Leu Asp Tyr Asn Val Asp Tyr Gly Ala Cys Val Thr Ala Ser
915 920 925

Thr Arg Phe Gly Asp Gly Asn Ala Arg Ser Ser Ile Ile Asn Phe Arg
930 935 940

Thr Pro Glu Gly Glu Pro Ser Asp Pro Pro Asn Asp Val His Tyr Val
945 950 955 960

Asn Leu Ser Ser Ser Ser Ile Ile Leu Phe Trp Thr Pro Pro Val Lys
965 970 975

Pro Asn Gly Ile Ile Gln Tyr Tyr Ser Val Tyr Tyr Gln Asn Thr Ser
980 985 990

Gly Thr Phe Val Gln Asn Phe Thr Leu Leu Gln Val Thr Lys Glu Ser
995 1000 1005

Asp Asn Val Thr Val Ser Ala Arg Ile Tyr Arg Leu Ala Ile Phe
1010 1015 1020

Ser Tyr Tyr Thr Phe Trp Leu Thr Ala Ser Thr Ser Val Gly Asn
1025 1030 1035

Gly Asn Lys Ser Ser Asp Ile Ile His Val Tyr Thr Asp Gln Asp
1040 1045 1050

Ile Pro Glu Gly Pro Val Gly Asn Leu Thr Phe Glu Ser Ile Ser
1055 1060 1065

Ser Thr Ala Ile His Val Ser Trp Glu Pro Pro Ser Gln Pro Asn
1070 1075 1080

Gly Leu Val Phe Tyr Tyr Leu Ser Leu Asn Leu Gln Gln Ser Pro
1085 1090 1095

Pro Arg His Met Ile Pro Pro Leu Val Thr Tyr Glu Asn Ser Ile
1100 1105 1110

Asp Phe Asp Asp Leu Glu Lys Tyr Thr Asp Tyr Ile Phe Lys Ile
1115 1120 1125

Thr Pro Ser Thr Glu Lys Gly Phe Ser Glu Thr Tyr Thr Thr Gln
1130 1135 1140

Leu His Ile Lys Thr Glu Glu Asp Val Pro Asp Thr Pro Pro Ile
1145 1150 1155

Ile Asn Thr Phe Lys Asn Leu Ser Ser Thr Ser Ile Leu Leu Ser
1160 1165 1170

Trp Asp Pro Pro Leu Lys Pro Asn Gly Ala Ile Leu Gly Tyr His
1175 1180 1185

Leu Thr Leu Gln Gly Pro His Ala Asn His Thr Phe Val Thr Ser
1190 1195 1200

Gly Asn His Ile Val Leu Glu Glu Leu Ser Pro Phe Thr Leu Tyr
1205 1210 1215

Ser Phe Phe Ala Ala Ala Arg Thr Met Lys Gly Leu Gly Pro Ser
1220 1225 1230

Ser Ile Leu Phe Phe Tyr Thr Asp Glu Ser Ala Pro Leu Ala Pro
1235 1240 1245

Pro Gln Asn Leu Thr Leu Ile Asn Tyr Thr Ser Asp Phe Val Trp
1250 1255 1260

Leu Thr Trp Ser Pro Ser Pro Leu Pro Gly Gly Ile Val Lys Val
1265 1270 1275

Tyr Ser Phe Lys Ile His Glu His Glu Thr Asp Thr Val Phe Tyr
1280 1285 1290

Lys Asn Ile Ser Gly Leu Gln Thr Asp Ala Lys Leu Glu Gly Leu
1295 1300 1305

Glu Pro Val Ser Thr Tyr Ser Val Ser Val Ser Ala Phe Thr Lys
1310 1315 1320

Val Gly Asn Gly Asn Gln Tyr Ser Asn Val Val Glu Phe Thr Thr
1325 1330 1335

Gln Glu Ser Val Pro Glu Ala Val Arg Asn Ile Glu Cys Val Ala
1340 1345 1350

Arg Asp Trp Gln Ser Val Ser Val Arg Trp Asp Pro Pro Arg Lys
1355 1360 1365

Thr Asn Gly Ile Ile Ile His Tyr Met Ile Thr Val Gly Gly Asn
1370 1375 1380

Ser Thr Lys Val Ser Pro Arg Asp Pro Thr Tyr Thr Phe Thr Lys
1385 1390 1395

Leu Leu Pro Asn Thr Ser Tyr Val Phe Glu Val Arg Ala Ser Thr
1400 1405 1410

Ser Ala Gly Glu Gly Asn Glu Ser Arg Cys Asp Ile Ser Thr Leu
1415 1420 1425

Pro Glu Thr Val Pro Ser Ala Pro Thr Asn Val Ala Phe Ser Asn
1430 1435 1440

Val Gln Ser Thr Ser Ala Thr Leu Thr Trp Thr Lys Pro Asp Thr
1445 1450 1455

Ile Phe Gly Tyr Phe Gln Asn Tyr Lys Ile Thr Thr Gln Leu Arg
1460 1465 1470

Ala Gln Lys Cys Arg Glu Trp Glu Pro Glu Glu Cys Ile Glu His
1475 1480 1485

Gln Lys Asp Gln Tyr Leu Tyr Glu Ala Asn Gln Thr Glu Glu Thr
1490 1495 1500

Val His Gly Leu Lys Lys Phe Arg Trp Tyr Arg Phe Gln Val Ala
1505 1510 1515

Ala Ser Thr Asn Val Gly Tyr Ser Asn Ala Ser Glu Trp Ile Ser
1520 1525 1530

Thr Gln Thr Leu Pro Gly Pro Pro Asp Gly Pro Pro Glu Asn Val
1535 1540 1545

His Val Val Ala Thr Ser Pro Phe Gly Ile Asn Ile Ser Trp Ser
1550 1555 1560

Glu Pro Ala Val Ile Thr Gly Pro Thr Phe Tyr Leu Ile Asp Val
1565 1570 1575

Lys Ser Val Asp Asp Asp Asp Phe Asn Ile Ser Phe Leu Lys Ser
1580 1585 1590

Asn Glu Glu Asn Lys Thr Thr Glu Ile Asn Asn Leu Glu Val Phe
1595 1600 1605

Thr Arg Tyr Ser Val Val Ile Thr Ala Phe Val Gly Asn Val Ser
1610 1615 1620

Arg Ala Tyr Thr Asp Gly Lys Ser Ser Ala Glu Val Ile Ile Thr
1625 1630 1635

Thr Leu Glu Ser Val Pro Lys Asp Pro Pro Asn Asn Met Thr Phe
1640 1645 1650

Gln Lys Ile Pro Asp Glu Val Thr Lys Phe Gln Leu Thr Phe Leu
1655 1660 1665

Pro Pro Ser Gln Pro Asn Gly Asn Ile Arg Val Tyr Gln Ala Leu
1670 1675 1680

Val Tyr Arg Glu Asp Asp Pro Thr Ala Val Gln Ile His Asn Phe
1685 1690 1695

Ser Ile Ile Gln Lys Thr Asp Thr Ser Ile Ile Ala Met Leu Glu
1700 1705 1710

Gly Leu Lys Gly Gly His Thr Tyr Asn Ile Ser Val Tyr Ala Ile
1715 1720 1725

Asn Ser Ala Gly Ala Gly Pro Lys Val Gln Met Arg Ile Thr Met
1730 1735 1740

Asp Ile Lys Ala Pro Ala Arg Pro Lys Ser Lys Pro Ile Pro Ile
1745 1750 1755

Arg Asp Ala Thr Gly Lys Leu Leu Val Thr Ser Thr Thr Ile Thr
1760 1765 1770

Ile Arg Met Pro Ile Cys Tyr Tyr Asn Asp Asp His Gly Pro Ile
1775 1780 1785

Arg Asn Val Gln Val Leu Val Ala Glu Thr Gly Ala Gln Gln Asp
1790 1795 1800

Gly Asn Val Thr Lys Trp Tyr Asp Ala Tyr Phe Asn Lys Ala Arg
1805 1810 1815

Pro Tyr Phe Thr Asn Glu Gly Phe Pro Asn Pro Pro Cys Ile Glu
1820 1825 1830

Gly Lys Thr Lys Phe Ser Gly Asn Glu Glu Ile Tyr Val Ile Gly
1835 1840 1845

Ala Asp Asn Ala Cys Met Ile Pro Gly Asn Glu Glu Lys Ile Cys
1850 1855 1860

Asn Gly Pro Leu Lys Pro Lys Lys Gln Tyr Leu Phe Lys Phe Arg
1865 1870 1875

Ala Thr Asn Val Met Gly Gln Phe Thr Asp Ser Glu Tyr Ser Asp
1880 1885 1890

Pro Ile Lys Thr Leu Gly Glu Gly Leu Ser Glu Arg Thr Val Glu
1895 1900 1905

Ile Ile Leu Ser Val Thr Leu Cys Ile Leu Ser Ile Ile Leu Leu
1910 1915 1920

Gly Thr Ala Ile Phe Ala Phe Val Arg Ile Arg Gln Lys Gln Lys
1925 1930 1935

Glu Gly Gly Thr Tyr Ser Pro Arg Asp Ala Glu Ile Ile Asp Thr
1940 1945 1950

Lys Phe Lys Leu Asp Gln Leu Ile Thr Val Ala Asp Leu Glu Leu
1955 1960 1965

Lys Asp Glu Arg Leu Thr Arg Leu Leu Ser Tyr Arg Lys Ser Ile
1970 1975 1980

Lys Pro Ile Ser Lys Lys Ser Phe Leu Gln His Val Glu Glu Leu
1985 1990 1995

Cys Thr Asn Ser Asn Leu Lys Phe Gln Glu Glu Phe Ser Glu Leu
2000 2005 2010

Pro Lys Phe Leu Gln Asp Leu Ser Ser Thr Asp Ala Asp Leu Pro
2015 2020 2025

Trp Asn Arg Ala Lys Asn Arg Phe Pro Asn Ile Lys Pro Tyr Asn
2030 2035 2040

Asn Asn Arg Val Lys Leu Ile Ala Asp Val Ser Leu Pro Gly Ser
2045 2050 2055

Asp Tyr Ile Asn Ala Ser Tyr Val Ser Gly Tyr Leu Cys Pro Asn
2060 2065 2070

Glu Phe Ile Ala Thr Gln Gly Pro Leu Pro Gly Thr Val Gly Asp
2075 2080 2085

Phe Trp Arg Met Val Trp Glu Thr Arg Thr Lys Thr Leu Val Met
2090 2095 2100

Leu Thr Gln Cys Phe Glu Lys Gly Arg Ile Arg Cys His Gln Tyr
2105 2110 2115

Trp Pro Glu Asp Asn Lys Pro Val Thr Val Phe Gly Asp Ile Val
2120 2125 2130

Ile Thr Lys Leu Met Glu Asp Ile Gln Ile Asp Trp Thr Ile Arg
2135 2140 2145

Asp Leu Lys Ile Glu Arg His Gly Asp Cys Met Thr Val Arg Gln
2150 2155 2160

Cys Asn Phe Thr Gly Trp Pro Glu His Gly Val Pro Glu Asn Thr
2165 2170 2175

Thr Pro Leu Ile His Phe Val Lys Leu Val Arg Thr Ser Arg Ala
2180 2185 2190

His Asp Thr Thr Pro Met Val Val His Cys Ser Ala Gly Val Gly
2195 2200 2205

Arg Thr Gly Val Phe Ile Ala Leu Asp His Leu Thr Gln His Ile
2210 2215 2220

Asn Asn His Asp Phe Val Asp Ile Tyr Gly Leu Val Ala Glu Leu
2225 2230 2235

Arg Ser Glu Arg Met Cys Met Val Gln Asn Leu Ala Gln Tyr Ile
2240 2245 2250

Phe Leu His Gln Cys Ile Leu Asp Leu Leu Ser Asn Lys Gly Gly
2255 2260 2265

His Gln Pro Val Cys Phe Val Asn Tyr Ser Thr Leu Gln Lys Met
2270 2275 2280

Asp Ser Leu Asp Ala Met Glu Gly Asp Val Glu Leu Glu Trp Glu
2285 2290 2295

Glu Thr Thr Met
2300

<210> 4

<211> 330

<212> PRT

<213> Homo sapiens

<400> 4

Met Asp Phe Leu Ile Ile Phe Leu Leu Leu Phe Ile Gly Thr Ser Glu
1 5 10 15

Thr Gln Val Asp Val Ser Asn Val Val Pro Gly Thr Arg Tyr Asp Ile
20 25 30

Thr Ile Ser Ser Ile Ser Thr Thr Tyr Thr Ser Pro Val Thr Arg Ile
35 40 45

Val Thr Thr Asn Val Thr Lys Pro Gly Pro Pro Val Phe Leu Ala Gly
50 55 60

Glu Arg Val Gly Ser Ala Gly Ile Leu Leu Ser Trp Asn Thr Pro Pro
65 70 75 80

Asn Pro Asn Gly Arg Ile Ile Ser Tyr Ile Val Lys Tyr Lys Glu Val
85 90 95

Cys Pro Trp Met Gln Thr Val Tyr Thr Gln Val Arg Ser Lys Pro Asp
100 105 110

Ser Leu Glu Val Leu Leu Thr Asn Leu Asn Pro Gly Thr Thr Tyr Glu
115 120 125

Ile Lys Val Ala Ala Glu Asn Ser Ala Gly Ile Gly Val Phe Ser Asp
130 135 140

Pro Phe Leu Phe Gln Thr Ala Glu Ser Ala Pro Gly Lys Val Val Asn
145 150 155 160

Leu Thr Val Glu Ala Tyr Asn Ala Ser Ala Val Lys Leu Ile Trp Tyr
165 170 175

Leu Pro Arg Gln Pro Asn Gly Lys Ile Thr Ser Phe Lys Ile Ser Val
180 185 190

Lys His Ala Arg Ser Gly Ile Val Val Lys Asp Val Ser Ile Arg Val
195 200 205

Glu Asp Ile Leu Thr Gly Lys Leu Pro Glu Ser Asn Glu Asn Ser Glu
210 215 220

Ser Phe Leu Trp Ser Thr Ala Ser Pro Ser Pro Thr Leu Gly Arg Val
225 230 235 240

Thr Pro Pro Ser Arg Thr Thr His Ser Ser Ser Thr Leu Thr Gln Asn
245 250 255

Glu Ile Ser Ser Val Trp Lys Glu Pro Ile Ser Phe Val Val Thr His
260 265 270

Leu Arg Pro Tyr Thr Thr Tyr Leu Phe Glu Val Ser Ala Ala Thr Thr
275 280 285

Glu Ala Gly Tyr Ile Asp Ser Thr Ile Val Arg Thr Pro Glu Ser Ala

290 295 300

Thr Ala Ala Gly Thr Gln Ala Asn Arg Val Trp Ser Arg Pro Leu Ala
305 310 315 320

Asn Ser Asn Arg Pro Ala Ala Glu Gly Pro
325 330

<210> 5

<211> 1959

<212> PRT

<213> Homo sapiens

<400> 5

Leu Leu Gly Gly Lys Leu Thr Asn Arg Lys Asp Ile His Thr Lys Asn
1 5 10 15

Pro Ser Val His His His His Gln Arg Pro Lys Val Asp Lys Thr Thr
20 25 30

Lys Met Gly Lys Lys Gln Ser Arg Lys Thr Gly Asn Ser Lys Lys Gln
35 40 45

Ser Thr Ser Pro Pro Pro Lys Asp Gly Arg Ala Val Phe Asp Leu Gln
50 55 60

Leu Ala Glu Val Glu Ser Thr Gln Val Arg Ile Thr Trp Lys Lys Pro
65 70 75 80

Arg Gln Pro Asn Gly Ile Ile Asn Gln Tyr Arg Val Lys Val Leu Val
85 90 95

Pro Glu Thr Gly Ile Ile Leu Glu Asn Thr Leu Leu Thr Gly Asn Asn
100 105 110

Glu Tyr Ile Asn Asp Pro Met Ala Pro Glu Ile Val Asn Ile Val Glu
115 120 125

Pro Met Val Gly Leu Tyr Glu Gly Ser Ala Glu Met Ser Ser Asp Leu
130 135 140

His Ser Leu Ala Thr Phe Ile Tyr Asn Ser His Pro Asp Lys Asn Phe
145 150 155 160

Pro Ala Arg Asn Arg Ala Glu Asp Gln Thr Ser Pro Val Val Thr Thr
165 170 175

Arg Asn Gln Tyr Ile Thr Asp Ile Ala Ala Glu Gln Leu Ser Tyr Val
180 185 190

Ile Arg Arg Leu Val Pro Phe Thr Glu His Met Ile Ser Val Ser Ala
195 200 205

Phe Thr Ile Met Gly Glu Gly Pro Pro Thr Val Leu Ser Val Arg Thr
210 215 220

Arg Gln Gln Val Pro Ser Ser Ile Lys Ile Ile Asn Tyr Lys Asn Ile
225 230 235 240

Ser Ser Ser Ser Ile Leu Leu Tyr Trp Asp Pro Pro Glu Tyr Pro Asn
245 250 255

Gly Lys Ile Thr His Tyr Thr Ile Tyr Ala Met Glu Leu Asp Thr Asn
260 265 270

Arg Ala Phe Gln Ile Thr Thr Ile Asp Asn Ser Phe Leu Ile Thr Gly
275 280 285

Leu Lys Lys Tyr Thr Lys Tyr Lys Met Arg Val Ala Ala Ser Thr His
290 295 300

Val Gly Glu Ser Ser Leu Ser Glu Glu Asn Asp Ile Phe Val Arg Thr
305 310 315 320

Ser Glu Asp Glu Pro Glu Ser Ser Pro Gln Asp Val Glu Val Ile Asp
325 330 335

Val Thr Ala Asp Glu Ile Arg Leu Lys Trp Ser Pro Pro Glu Lys Pro
340 345 350

Asn Gly Ile Ile Ile Ala Tyr Glu Val Leu Tyr Lys Asn Ile Asp Thr
355 360 365

Leu Tyr Met Lys Asn Thr Ser Thr Thr Asp Ile Ile Leu Arg Asn Leu
370 375 380

Arg Pro His Thr Leu Tyr Asn Ile Ser Val Arg Ser Tyr Thr Arg Phe
385 390 395 400

Gly His Gly Asn Gln Val Ser Ser Leu Leu Ser Val Arg Thr Ser Glu
405 410 415

Thr Val Pro Asp Ser Ala Pro Glu Asn Ile Thr Tyr Lys Asn Ile Ser
420 425 430

Ser Gly Glu Ile Glu Leu Ser Phe Leu Pro Pro Ser Ser Pro Asn Gly
435 440 445

Ile Ile Gln Lys Tyr Thr Ile Tyr Leu Lys Arg Ser Asn Gly Asn Glu
450 455 460

Glu Arg Thr Ile Asn Thr Thr Ser Leu Thr Gln Asn Ile Lys Val Leu
465 470 475 480

Lys Lys Tyr Thr Gln Tyr Ile Ile Glu Val Ser Ala Ser Thr Leu Lys
485 490 495

Gly Glu Gly Val Arg Ser Ala Pro Ile Ser Ile Leu Thr Glu Glu Asp
500 505 510

Ala Pro Asp Ser Pro Pro Gln Asp Phe Ser Val Lys Gln Leu Ser Gly
515 520 525

Val Thr Val Lys Leu Ser Trp Gln Pro Pro Leu Glu Pro Asn Gly Ile
530 535 540

Ile Leu Tyr Tyr Thr Val Tyr Val Trp Asn Arg Ser Ser Leu Lys Thr
545 550 555 560

Ile Asn Val Thr Glu Thr Ser Leu Glu Leu Ser Asp Leu Asp Tyr Asn
565 570 575

Val Glu Tyr Ser Ala Tyr Val Thr Ala Ser Thr Arg Phe Gly Asp Gly
580 585 590

Lys Thr Arg Ser Asn Ile Ile Ser Phe Gln Thr Pro Glu Gly Ala Pro
595 600 605

Ser Asp Pro Pro Lys Asp Val Tyr Tyr Ala Asn Leu Ser Ser Ser Ser
610 615 620

```
Ile Ile Leu Phe Trp Thr Pro Pro Ser Lys Pro Asn Gly Ile Ile Gln
625                 630                 635                 640

Tyr Tyr Ser Val Tyr Tyr Arg Asn Thr Ser Gly Thr Phe Met Gln Asn
                645                 650                 655

Phe Thr Leu His Glu Val Thr Asn Asp Phe Asp Asn Met Thr Val Ser
            660                 665                 670

Thr Ile Ile Asp Lys Leu Thr Ile Phe Ser Tyr Tyr Thr Phe Trp Leu
        675                 680                 685

Thr Ala Ser Thr Ser Val Gly Asn Gly Asn Lys Ser Ser Asp Ile Ile
    690                 695                 700

Glu Val Tyr Thr Asp Gln Asp Ile Pro Glu Gly Phe Val Gly Asn Leu
705                 710                 715                 720

Thr Tyr Glu Ser Ile Ser Ser Thr Ala Ile Asn Val Ser Trp Val Pro
                725                 730                 735

Pro Ala Gln Pro Asn Gly Leu Val Phe Tyr Tyr Val Ser Leu Ile Leu
            740                 745                 750

Gln Gln Thr Pro Arg His Val Arg Pro Pro Leu Val Thr Tyr Glu Arg
        755                 760                 765

Ser Ile Tyr Phe Asp Asn Leu Glu Lys Tyr Thr Asp Tyr Ile Leu Lys
    770                 775                 780

Ile Thr Pro Ser Thr Glu Lys Gly Phe Ser Asp Thr Tyr Thr Ala Gln
785                 790                 795                 800

Leu Tyr Ile Lys Thr Glu Glu Asp Val Pro Glu Thr Ser Pro Ile Ile
                805                 810                 815

Asn Thr Phe Lys Asn Leu Ser Ser Thr Ser Val Leu Leu Ser Trp Asp
            820                 825                 830

Pro Pro Val Lys Pro Asn Gly Ala Ile Ile Ser Tyr Asp Leu Thr Leu
        835                 840                 845
```

Gln Gly Pro Asn Glu Asn Tyr Ser Phe Ile Thr Ser Asp Asn Tyr Ile
850 855 860

Ile Leu Glu Glu Leu Ser Pro Phe Thr Leu Tyr Ser Phe Phe Ala Ala
865 870 875 880

Ala Arg Thr Arg Lys Gly Leu Gly Pro Ser Ser Ile Leu Phe Phe Tyr
885 890 895

Thr Asp Glu Ser Val Pro Leu Ala Pro Pro Gln Asn Leu Thr Leu Ile
900 905 910

Asn Cys Thr Ser Asp Phe Val Trp Leu Lys Trp Ser Pro Ser Pro Leu
915 920 925

Pro Gly Gly Ile Val Lys Val Tyr Ser Phe Lys Ile His Glu His Glu
930 935 940

Thr Asp Thr Ile Tyr Tyr Lys Asn Ile Ser Gly Phe Lys Thr Glu Ala
945 950 955 960

Lys Leu Val Gly Leu Glu Pro Val Ser Thr Tyr Ser Ile Arg Val Ser
965 970 975

Ala Phe Thr Lys Val Gly Asn Gly Asn Gln Phe Ser Asn Val Val Lys
980 985 990

Phe Thr Thr Gln Glu Ser Val Pro Asp Val Val Gln Asn Met Gln Cys
995 1000 1005

Met Ala Thr Ser Trp Gln Ser Val Leu Val Lys Trp Asp Pro Pro
1010 1015 1020

Lys Lys Ala Asn Gly Ile Ile Thr Gln Tyr Met Val Thr Val Glu
1025 1030 1035

Arg Asn Ser Thr Lys Val Ser Pro Gln Asp His Met Tyr Thr Phe
1040 1045 1050

Ile Lys Leu Leu Ala Asn Thr Ser Tyr Val Phe Lys Val Arg Ala
1055 1060 1065

Ser Thr Ser Ala Gly Glu Gly Asp Glu Ser Thr Cys His Val Ser
1070 1075 1080

```
Thr Leu  Pro Glu Thr Val Pro  Ser Val Pro Thr Asn  Ile Ala Phe
    1085                 1090                 1095

Ser Asp  Val Gln Ser Thr Ser  Ala Thr Leu Thr Trp  Ile Arg Pro
    1100                 1105                 1110

Asp Thr  Ile Leu Gly Tyr Phe  Gln Asn Tyr Lys Ile  Thr Thr Gln
    1115                 1120                 1125

Leu Arg  Ala Gln Lys Cys Lys  Glu Trp Glu Ser Glu  Glu Cys Val
    1130                 1135                 1140

Glu Tyr  Gln Lys Ile Gln Tyr  Leu Tyr Glu Ala His  Leu Thr Glu
    1145                 1150                 1155

Glu Thr  Val Tyr Gly Leu Lys  Lys Phe Arg Trp Tyr  Arg Phe Gln
    1160                 1165                 1170

Val Ala  Ala Ser Thr Asn Ala  Gly Tyr Gly Asn Ala  Ser Asn Trp
    1175                 1180                 1185

Ile Ser  Thr Lys Thr Leu Pro  Gly Pro Pro Asp Gly  Pro Pro Glu
    1190                 1195                 1200

Asn Val  His Val Val Ala Thr  Ser Pro Phe Ser Ile  Ser Ile Ser
    1205                 1210                 1215

Trp Ser  Glu Pro Ala Val Ile  Thr Gly Pro Thr Cys  Tyr Leu Ile
    1220                 1225                 1230

Asp Val  Lys Ser Val Asp Asn  Asp Glu Phe Asn Ile  Ser Phe Ile
    1235                 1240                 1245

Lys Ser  Asn Glu Glu Asn Lys  Thr Ile Glu Ile Lys  Asp Leu Glu
    1250                 1255                 1260

Ile Phe  Thr Arg Tyr Ser Val  Val Ile Thr Ala Phe  Thr Gly Asn
    1265                 1270                 1275

Ile Ser  Ala Ala Tyr Val Glu  Gly Lys Ser Ser Ala  Glu Met Ile
    1280                 1285                 1290

Val Thr  Thr Leu Glu Ser Ala  Pro Lys Asp Pro Pro  Asn Asn Met
    1295                 1300                 1305
```

Thr Phe Gln Lys Ile Pro Asp Glu Val Thr Lys Phe Gln Leu Thr
1310 1315 1320

Phe Leu Pro Pro Ser Gln Pro Asn Gly Asn Ile Gln Val Tyr Gln
1325 1330 1335

Ala Leu Val Tyr Arg Glu Asp Asp Pro Thr Ala Val Gln Ile His
1340 1345 1350

Asn Leu Ser Ile Ile Gln Lys Thr Asn Thr Phe Val Ile Ala Met
1355 1360 1365

Leu Glu Gly Leu Lys Gly Gly His Thr Tyr Asn Ile Ser Val Tyr
1370 1375 1380

Ala Val Asn Ser Ala Gly Ala Gly Pro Lys Val Pro Met Arg Ile
1385 1390 1395

Thr Met Asp Ile Lys Ala Pro Ala Arg Pro Lys Thr Lys Pro Thr
1400 1405 1410

Pro Ile Tyr Asp Ala Thr Gly Lys Leu Leu Val Thr Ser Thr Thr
1415 1420 1425

Ile Thr Ile Arg Met Pro Ile Cys Tyr Tyr Ser Asp Asp His Gly
1430 1435 1440

Pro Ile Lys Asn Val Gln Val Leu Val Thr Glu Thr Gly Ala Gln
1445 1450 1455

His Asp Gly Asn Val Thr Lys Trp Tyr Asp Ala Tyr Phe Asn Lys
1460 1465 1470

Ala Arg Pro Tyr Phe Thr Asn Glu Gly Phe Pro Asn Pro Pro Cys
1475 1480 1485

Thr Glu Gly Lys Thr Lys Phe Ser Gly Asn Glu Glu Ile Tyr Ile
1490 1495 1500

Ile Gly Ala Asp Asn Ala Cys Met Ile Pro Gly Asn Glu Asp Lys
1505 1510 1515

Ile Cys Asn Gly Pro Leu Lys Pro Lys Lys Gln Tyr Leu Phe Lys
1520 1525 1530

Phe Arg Ala Thr Asn Ile Met Gly Gln Phe Thr Asp Ser Asp Tyr
1535 1540 1545

Ser Asp Pro Val Lys Thr Leu Gly Glu Gly Leu Ser Glu Arg Thr
1550 1555 1560

Val Glu Ile Ile Leu Ser Val Thr Leu Cys Ile Leu Ser Ile Ile
1565 1570 1575

Leu Leu Gly Thr Ala Ile Phe Ala Phe Ala Arg Ile Arg Gln Lys
1580 1585 1590

Gln Lys Glu Gly Gly Thr Tyr Ser Pro Gln Asp Ala Glu Ile Ile
1595 1600 1605

Asp Thr Lys Leu Lys Leu Asp Gln Leu Ile Thr Val Ala Asp Leu
1610 1615 1620

Glu Leu Lys Asp Glu Arg Leu Thr Arg Leu Leu Ser Tyr Arg Lys
1625 1630 1635

Ser Ile Lys Pro Ile Ser Lys Lys Ser Phe Leu Gln His Val Glu
1640 1645 1650

Glu Leu Cys Thr Asn Asn Asn Leu Lys Phe Gln Glu Glu Phe Ser
1655 1660 1665

Glu Leu Pro Lys Phe Leu Gln Asp Leu Ser Ser Thr Asp Ala Asp
1670 1675 1680

Leu Pro Trp Asn Arg Ala Lys Asn Arg Phe Pro Asn Ile Lys Pro
1685 1690 1695

Tyr Asn Asn Asn Arg Val Lys Leu Ile Ala Asp Ala Ser Val Pro
1700 1705 1710

Gly Ser Asp Tyr Ile Asn Ala Ser Tyr Ile Ser Gly Tyr Leu Cys
1715 1720 1725

Pro Asn Glu Phe Ile Ala Thr Gln Gly Pro Leu Pro Gly Thr Val
1730 1735 1740

Gly Asp Phe Trp Lys Met Val Trp Glu Thr Arg Ala Lys Thr Leu
1745 1750 1755

Val Met Leu Thr Gln Cys Phe Glu Lys Gly Arg Ile Arg Cys His
1760 1765 1770

Gln Tyr Trp Pro Glu Asp Asn Lys Pro Val Thr Val Phe Gly Asp
1775 1780 1785

Ile Val Ile Thr Lys Leu Met Glu Asp Val Gln Ile Asp Trp Thr
1790 1795 1800

Ile Arg Asp Leu Lys Ile Glu Arg His Gly Asp Cys Met Thr Val
1805 1810 1815

Arg Gln Cys Asn Phe Thr Ala Trp Pro Glu His Gly Val Pro Glu
1820 1825 1830

Asn Ser Ala Pro Leu Ile His Phe Val Lys Leu Val Arg Ala Ser
1835 1840 1845

Arg Ala His Asp Thr Thr Pro Met Ile Val His Cys Ser Ala Gly
1850 1855 1860

Val Gly Arg Thr Gly Val Phe Ile Ala Leu Asp His Leu Thr Gln
1865 1870 1875

His Ile Asn Asp His Asp Phe Val Asp Ile Tyr Gly Leu Val Ala
1880 1885 1890

Glu Leu Arg Ser Glu Arg Met Cys Met Val Gln Asn Leu Ala Gln
1895 1900 1905

Tyr Ile Phe Leu His Gln Cys Ile Leu Asp Leu Leu Ser Asn Lys
1910 1915 1920

Gly Ser Asn Gln Pro Ile Cys Phe Val Asn Tyr Ser Ala Leu Gln
1925 1930 1935

Lys Met Asp Ser Leu Asp Ala Met Glu Gly Asp Val Glu Leu Glu
1940 1945 1950

Trp Glu Glu Thr Thr Met
1955

<210> 6

<211> 990

<212> DNA

<213> Homo sapiens

<400> 6
atggattttc ttatcatttt tcttttactt tttattggga cttcagagac acaggttgat 60

gtttccaatg tcgttcctgg tactaggtac gatataacca tctcttcaat ttctacaaca 120

tacacctcac ctgttactag aatagtgaca acaaatgtaa caaaaccagg gcctccagtc 180

ttcctagccg gggaaagagt cggatctgct gggattcttc tgtcttggaa tacaccacct 240

aatccaaatg gaaggattat atcttacatt gtcaaatata aggaagtttg tccgtggatg 300

caaacagtat atacacaagt cagatcaaag ccagacagtc tggaagttct tcttactaat 360

cttaatcctg gaacaacata tgaaattaag gttgctgctg aaaacagtgc tggcattgga 420

gtgtttagtg atccatttct cttccaaact gcagaaagtg ctccaggaaa agtggtgaat 480

ctcacagttg aggcctacaa cgcttcagca gttaagctga tttggtattt acctcggcaa 540

ccaaatggca aaattaccag cttcaagatt agtgtcaagc atgccagaag tgggatagta 600

gtgaaagatg tctcaatcag agtagaggac attttgactg ggaaattgcc agaaagcaat 660

gagaatagtg aatctttttt atggagtaca gccagccctt ctccaaccct tggtagagtt 720

acacctccat cgcgtaccac acattcatca agcacgttga cacagaatga gatcagctct 780

gtgtggaaag agcctatcag ttttgtagtg acacacttga gaccttatac aacatatctt 840

tttgaagttt cagctgctac aactgaagca ggttatattg atagtacgat tgtcagaaca 900

ccagaatcag ccaccgctgc tggtacccag gctaacaggg tctggagtag acctctagca 960

aactccaaca gacctgcagc ggagggtcct 990

<210> 7

<211> 5877

<212> DNA

<213> Homo sapiens

<400> 7

```
ctgttaggag gaaaactaac aaacagaaag gacatccaca ccaaaaaccc atctgtacat      60
caccatcatc aaagaccaaa agtagataaa accacaaaga tggggaaaaa acagagcaga     120
aaaactggaa actctaaaaa gcagagcacc tctcctcctc caaaggatgg cagggcagtg     180
tttgatttac aacttgcaga ggtagaatcc acgcaagtaa gaattacttg gaagaaacca     240
cgacaaccaa atggaattat taaccaatac cgagtgaaag tgctagttcc agagacagga     300
ataattttgg aaaatacttt gctcactgga aataatgagt atataaatga ccccatggct     360
ccagaaattg tgaacatagt agagccaatg gtaggattat atgagggttc agcagagatg     420
tcgtctgacc ttcactcact tgctacattt atatataaca gccatccaga taaaaacttt     480
cctgcaagga atagagctga agaccagact tcaccagttg taactacaag gaatcagtat     540
attactgaca ttgcagctga acagctgtct tatgttatca ggagacttgt acctttcact     600
gagcacatga ttagtgtatc tgctttcacc atcatgggag aaggaccacc aacagttctc     660
agtgttagga cacgtcagca agtgccaagc tccattaaaa ttataaacta taaaaatatt     720
agttcttcat ctattttgtt atattgggat cctccagaat atcccaatgg aaaaataact     780
cactatacga tttatgcaat ggaattggat acaaacagag cattccagat aactaccata     840
gataacagct ttctcataac agggttaaag aaatacacaa aatacaaaat gagagtggca     900
gcctcaaccc acgttggaga aagttctttg tctgaagaaa atgacatctt tgtgagaact     960
tcagaagatg aaccggaatc atcacctcaa gatgtcgaag taattgatgt taccgcagat    1020
gaaataaggt tgaagtggtc accacccgaa aagcccaatg ggatcattat tgcttatgaa    1080
gtgctatata aaaatataga tactttatat atgaagaaca catcaacaac agacataata    1140
ttaaggaact taagacctca caccctctat aacatttctg taaggtctta caccagattt    1200
ggtcatggca atcaggtatc ttctttactc tctgtaagga cttcggagac tgtgcctgat    1260
agtgcaccag aaaatatcac ttacaaaaat atttcttctg gagagattga gctatcattc    1320
cttcccccaa gtagtcccaa tggaatcata caaaaatata caatttatct caagagaagt    1380
aatggaaatg aggaaagaac tataaataca acctctttaa cccaaaacat taaagtactg    1440
aagaaatata cccaatatat cattgaggtg tctgctagta cactcaaagg tgaaggagtt    1500
cggagtgctc ccataagtat actgacggag gaagatgctc ctgattctcc ccctcaagac    1560
ttctctgtaa aacagttgtc tggtgtcacg gtgaagttgt catggcaacc accccctggag    1620
ccaaatggaa ttatccttta ttacacagtt tatgtctgga atagatcatc attaaaaact    1680
```

attaatgtca ctgaaacatc attggagtta tcagatttgg attataatgt tgaatacagt 1740
gcttatgtaa cagctagcac cagatttggt gatgggaaaa caagaagcaa tatcattagc 1800
tttcaaacac cagagggagc accaagcgat cctcccaaag atgtttatta tgcaaacctc 1860
agttcttcat caataattct tttctggaca cctccttcaa aacctaatgg gattatacaa 1920
tattactctg tttattacag aaatacttca ggtactttta tgcagaattt tacactccat 1980
gaagtaacca atgactttga caatatgact gtatccacaa ttatagataa actgacaata 2040
ttcagctact atacattttg gttaacagca agtacttcag ttggaaatgg gaataaaagc 2100
agtgacatca ttgaagtata cacagatcaa gacatacctg aagggtttgt tggaaacctg 2160
acttacgaat ccatttcgtc aactgcaata aatgtaagct gggtcccacc ggctcaacca 2220
aacggtctag tcttctacta tgtttcactg atcttacagc agactcctcg ccatgtgaga 2280
ccacctcttg ttacatatga gagaagcata tattttgata atctggaaaa atacactgat 2340
tatatattaa aaattactcc atcaacagaa aagggattct ctgataccta tactgcccag 2400
ctatacatca agactgaaga agatgtccca gaaacttcac caataatcaa cacttttaaa 2460
aacctttcct ctacctcagt tctcttatca tgggatcccc cagtaaagcc aaatggtgca 2520
ataataagtt atgatttaac tttacaagga ccaaatgaaa attattcttt cattacttct 2580
gataattaca taatattgga agagctttca ccatttacat tatatagctt ttttgctgcc 2640
gcaagaacta gaaaaggact tggtccttcc agtattcttt tcttttacac agatgagtca 2700
gtgccgttag cacctccaca aaatttgact ttaatcaact gtacttcaga ctttgtatgg 2760
ctgaaatgga gcccaagtcc tcttccaggt ggtattgtta aagtatatag ttttaaaatt 2820
catgaacatg aaactgacac tatatattat aagaatatat caggatttaa aactgaagcc 2880
aaacttgttg gactggaacc agtcagcacc tactctatcc gtgtatctgc gttcaccaaa 2940
gttggaaatg gcaatcaatt tagtaatgta gtaaaattca caacccaaga atcagttcca 3000
gatgtcgtgc agaatatgca gtgcatggca actagctggc agtcagtttt agtgaaatgg 3060
gatccaccca aaaaggcaaa tggaataata acgcagtata tggtaacagt tgaaaggaat 3120
tctacaaaag tttctcccca agatcacatg tacactttca taaagcttct tgccaatacc 3180
tcatatgtct ttaaagtaag agcttcaacc tcagctggtg aaggtgatga aagcacatgc 3240
catgtcagca cactacctga aacagttccc agtgttccca caaatattgc tttttctgat 3300
gttcagtcaa ctagtgcaac attgacatgg ataagacctg acactatcct tggctacttt 3360
caaaattaca aaattaccac tcaacttcgt gctcaaaaat gcaaagaatg ggaatccgaa 3420
gaatgtgttg aatatcaaaa aattcaatac ctctatgaag ctcacttaac tgaagagaca 3480

gtatatggat taaagaaatt tagatggtat agattccaag tggctgccag caccaatgct 3540
ggctatggca atgcttcaaa ctggatttct acaaaaactc tgcctggccc tccagatggt 3600
cctcctgaaa atgttcatgt agtagcaaca tcacctttta gcatcagcat aagctggagt 3660
gaacctgctg tcattactgg accaacatgt tatctgattg atgtcaaatc ggtagataat 3720
gatgaattta atatatcctt catcaagtca aatgaagaaa ataaaaccat agaaattaaa 3780
gatttagaaa tattcacaag gtattctgta gtgatcactg catttactgg gaacattagt 3840
gctgcatatg tagaagggaa gtcaagtgct gaaatgattg ttactacttt agaatcagcc 3900
ccaaaggacc cacctaacaa catgacattt cagaagatac cagatgaagt tacaaaattt 3960
caattaacgt tccttcctcc ttctcaacct aatggaaata tccaagtata tcaagctctg 4020
gtttaccgag aagatgatcc tactgctgtc cagattcaca acctcagtat tatacagaaa 4080
accaacacat tcgtcattgc aatgctagaa ggactaaaag gtggacatac atacaatatc 4140
agtgtttacg cagtcaatag tgctggtgca ggtccaaagg ttccgatgag aataaccatg 4200
gatatcaaag ctccagcacg accaaaaacc aaaccaaccc ctatttatga tgccacagga 4260
aaactgcttg tgacttcaac aacaattaca atcagaatgc caatatgtta ctacagtgat 4320
gatcatggac caataaaaaa tgtacaagtg cttgtgacag aaacaggagc tcagcatgat 4380
ggaaatgtaa caaagtggta tgatgcatat tttaataaag caaggccata ttttacaaat 4440
gaaggctttc ctaaccctcc atgtacagaa ggaaagacaa agtttagtgg caatgaagaa 4500
atctacatca taggtgctga taatgcatgc atgattcctg gcaatgaaga caaaatttgc 4560
aatggaccac tgaaaccaaa aaagcaatac ttatttaaat ttagagctac aaatattatg 4620
ggacaattta ctgactctga ttattctgac cctgttaaga ctttagggga aggactttca 4680
gaaagaaccg tagagatcat tctttccgtc actttgtgta tcctttcaat aattctcctt 4740
ggaacagcta tttttgcatt tgcaagaatt cgacagaagc agaaagaagg tggcacatac 4800
tctcctcagg atgcagaaat tattgacact aaattgaagc tggatcagct catcacagtg 4860
gcagacctgg aactgaagga cgagagatta acgcggttac ttagttatag aaaatccatc 4920
aagccaataa gcaagaaatc cttcctgcaa catgttgaag agctttgcac aaacaacaac 4980
ctaaagtttc aagaagaatt ttcggaatta ccaaaatttc ttcaggatct ttcttcaact 5040
gatgctgatc tgccttggaa tagagcaaaa aaccgcttcc caaacataaa accatataat 5100
aataacagag taaagctgat agctgacgct agtgttccag gttcggatta tattaatgcc 5160
agctatattt ctggttattt atgtccaaat gaatttattg ctactcaagg tccactacca 5220

ggaacagttg gagatttttg gaaaatggtg tgggaaacca gagcaaaaac attagtaatg 5280

ctaacacagt gttttgaaaa aggacggatc agatgccatc agtattggcc agaggacaac 5340

aagccagtta ctgtctttgg agatatagtg attacaaagc taatggagga tgttcaaata 5400

gattggacta tcagggatct gaaaattgaa aggcatgggg attgcatgac tgttcgacag 5460

tgtaacttta ctgcctggcc agagcatggg gttcctgaga acagcgcccc tctaattcac 5520

tttgtgaagt tggttcgagc aagcagggca catgacacca cacctatgat tgttcactgc 5580

agtgctggag ttggaagaac tggagttttt attgctctgg accatttaac acaacatata 5640

aatgaccatg attttgtgga tatatatgga ctagtagctg aactgagaag tgaaagaatg 5700

tgcatggtgc agaatctggc acagtatatc tttttacacc agtgcattct ggatctctta 5760

tcaaataagg gaagtaatca gcccatctgt tttgttaact attcagcact tcagaagatg 5820

gactctttgg acgccatgga aggtgatgtt gagcttgaat gggaagaaac cactatg 5877

<210> 8

<211> 412

<212> DNA

<213> Homo sapiens

<400> 8
tttactgctt aaaatctagg tgggtgtgag gccatgactg tgaacgttgt cagtcactga 60

ctgtattcag ctgtaaaggc ttgacaaggt gagaaaaacc tcccacactt ttaaaatgaa 120

tgaaatgttt gaacactaag tagtcataca caaaataatt accttaattt catatgttgt 180

tccaggatta agattagtaa gaagaacttc cagactgtct ggctttgatc tgacttgtgt 240

atatactgtt tgcatccacg gacaaacttc cttatatttg acaatgtaag atataatcct 300

tccatttgga ttaggtggtg tattccaaga cagaagaatc ccagcagatc cgactctttc 360

cccggctagg aagactggag gccctggttc tatttttaaa aaatccattt tg 412

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/11473

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N9/16 C12N15/55 C07K16/40 C12Q1/42 C12Q1/68
G01N33/573 A61K38/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, EPO-Internal, BIOSIS, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 21800 A (MAX PLANCK GESELLSCHAFT) 29 September 1994 (1994-09-29) cited in the application SEQ ID NOs:8, 10, 14, 21 and 22 abstract; claims 1-29; examples 10-14	1-17
X	WO 01 75075 A (BAYER AG; XIAO YONGHONG (US)) 11 October 2001 (2001-10-11) abstract; claims 2-17; examples 1-5	1-17
X	WO 01 20021 A (BOEHMER FRANK; HERRLICH PETER (DE); KARLSRUHE FORSCHZENT (DE)) 22 March 2001 (2001-03-22) abstract; claims 1-46	10-17

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 February 2003	24/02/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Sirim, P

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/11473

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WRIGHT MATTHEW B ET AL: "Proliferating and migrating mesangial cells responding to injury express a novel receptor protein-tyrosine phosphatase in experimental mesangial proliferative glomerulonephritis." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 37, pages 23929-23937, XP002229880 ISSN: 0021-9258 page 23937; figures 7,8 ---	1-17
A	ZHANG Z-Y: "Protein tyrosine phosphatases: Prospects for therapeutics" CURRENT OPINION IN CHEMICAL BIOLOGY, vol. 5, no. 4, 1 August 2001 (2001-08-01), pages 416-423, XP001039802 ISSN: 1367-5931 the whole document ---	1-17
A	OSTMAN ARNE ET AL: "Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatases" TRENDS IN CELL BIOLOGY, vol. 11, no. 6, June 2001 (2001-06), pages 258-266, XP002210401 ISSN: 0962-8924 abstract; figures 1,2; table 1 ---	1-17
P,X	WO 02 42435 A (BAYER AG; ZHU ZHIMIN (US)) 30 May 2002 (2002-05-30) abstract; claims 1-71; examples 1-6 -----	1-17

INTERNATIONAL SEARCH REPORT	Inter nal application No. PCT/EP 02/11473

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 13 to 17 (all in part)
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 13 to 17 (all in part)

The subject-matter of claims 13 and 14 which relates to agents defined by reference to a desirable property, namely their ability to modulate the activity the polypeptide of present application, or to methods involving said agents, was found to be ambiguous and vague as well as not sufficiently disclosed and supported (Art. 5 and 6 PCT).

Consequently, these claims have only been searched insofar as they relate to those agents which have been disclosed on page 57 of present application (e.g. antisense nucleic acid molecules)

Similarly, also claims 15 to 17 as far as they refer to claims 13 and 14 have only been partially searched.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/11473

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9421800	A	29-09-1994	US	5585233 A	17-12-1996
			AU	6429394 A	11-10-1994
			CA	2158939 A1	29-09-1994
			WO	9421800 A2	29-09-1994
			EP	0690918 A1	10-01-1996
			JP	8508877 T	24-09-1996
			US	6492495 B1	10-12-2002
			US	5952212 A	14-09-1999
WO 0175075	A	11-10-2001	AU	5223301 A	15-10-2001
			WO	0175075 A2	11-10-2001
WO 0120021	A	22-03-2001	DE	10035472 A1	15-03-2001
			WO	0120021 A2	22-03-2001
WO 0242435	A	30-05-2002	AU	3320102 A	03-06-2002
			WO	0242435 A2	30-05-2002